Brain tumor enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and generative artificial intelligence

Moya-Sáez, Elisa; Luis-García, Rodrigo de; Nunez-Gonzalez, Laura; Alberola-López, Carlos; Hernández-Tamames, Juan Antonio

DOI
10.21037/qims-24-721

Publication date
2025

Document Version
Final published version

Published in
Quantitative Imaging in Medicine and Surgery

Citation (APA)
Moya-Sáez, E., Luis-García, R. D., Nunez-Gonzalez, L., Alberola-López, C., & Hernández-Tamames, J. A. (2025). Brain tumor enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and generative artificial intelligence. *Quantitative Imaging in Medicine and Surgery*, *15*(1), 42-54. https://doi.org/10.21037/qims-24-721

Original Article

Brain tumor enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and generative artificial intelligence

Elisa Moya-Sáez[1]^, Rodrigo de Luis-García[1]^, Laura Nunez-Gonzalez[2], Carlos Alberola-López[1]^, Juan Antonio Hernández-Tamames[2,3]

[1]Image Processing Lab, University of Valladolid, Valladolid, Spain; [2]Radiology and Nuclear Medicine Department, Erasmus MC, Rotterdam, The Netherlands; [3]Imaging Physics Department, TU Delft, Delft, The Netherlands

Contributions: (I) Conception and design: All authors; (II) Administrative support: E Moya-Sáez, R de Luis-García, C Alberola-López, JA Hernández-Tamames; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Carlos Alberola-López, PhD. Image Processing Lab, University of Valladolid, Paseo Belén 15 47011 Valladolid, Spain. Email: carlos.alberola@uva.es.

Background: Gadolinium-based contrast agents (GBCAs) are usually employed for glioma diagnosis. However, GBCAs raise safety concerns, lead to patient discomfort and increase costs. Parametric maps offer a potential solution by enabling quantification of subtle tissue changes without GBCAs, but they are not commonly used in clinical practice due to the need for specifically targeted sequences. This work proposes to predict post-contrast T1-weighted enhancement without GBCAs from pre-contrast conventional weighted images through synthetic parametric maps computed with generative artificial intelligence (deep learning).

Methods: In this retrospective study, three datasets have been employed: (I) a proprietary dataset with 15 glioma patients (hereafter, GLIOMA dataset); (II) relaxometry maps from 5 healthy volunteers; and (III) UPenn-GBM, a public dataset with 493 glioblastoma patients. A deep learning method for synthesizing parametric maps from only two conventional weighted images is proposed. Particularly, we synthesize longitudinal relaxation time (T1), transversal relaxation time (T2), and proton density (PD) maps. The deep learning method is trained in a supervised manner with the GLIOMA dataset, which comprises weighted images and parametric maps obtained with magnetic resonance image compilation (MAGiC). Thus, MAGiC maps were used as references for the training. For testing, a leave-one-out scheme is followed. Finally, the synthesized maps are employed to predict T1-weighted enhancement without GBCAs. Our results are compared with those obtained by MAGiC; specifically, both the maps obtained with MAGiC and the synthesized maps are used to distinguish between healthy and abnormal tissue (ABN) and, particularly, tissues with and without T1-weighted enhancement. The generalization capability of the method was also tested on two additional datasets (healthy volunteers and the UPenn-GBM).

Results: Parametric maps synthesized with deep learning obtained similar performance compared to MAGiC for discriminating normal from ABN (sensitivities: 88.37% *vs.* 89.35%) and tissue with and without T1-weighted enhancement (sensitivities: 93.26% *vs.* 87.29%) on the GLIOMA dataset. These values were comparable to those obtained on UPenn-GBM (sensitivities of 91.23% and 81.04% for each classification).

Conclusions: Our results suggest the feasibility to predict T1-weighted-enhanced tissues from pre-contrast conventional weighted images using deep learning for the synthesis of parametric maps.

---

^ ORCID: Elisa Moya-Sáez, 0000-0003-4945-3193; Rodrigo de Luis-García, 0000-0001-5023-6490; Carlos Alberola-López, 0000-0003-3684-0055.

Quant Imaging Med Surg 2025;15(1):42-54 | https://dx.doi.org/10.21037/qims-24-721

Keywords: Enhancement prediction; synthetic magnetic resonance imaging (synthetic MRI); deep learning (DL); brain tumors; gadolinium-based contrast agents (GBCAs)

Submitted Apr 10, 2024. Accepted for publication Jul 22, 2024. Published online Dec 23, 2024.
doi: 10.21037/qims-24-721
View this article at: https://dx.doi.org/10.21037/qims-24-721

Introduction

Malignant gliomas are the most common type of primary brain tumors in adults with an incidence of approximately 5 per 100,000 people per year (1). Standard assessment with magnetic resonance imaging (MRI) includes T1-weighted (pre-T1w), T2-weighted (T2w), T2-weighted-fluid-attenuated inversion recovery (T2w-FLAIR), and T1w after the injection of a contrast agent (post-T1w) (2). Gadolinium-based contrast agents (GBCAs) are customary due to their ability to reveal the impairment in the blood-brain-barrier (BBB), associated with aggressive tumor behavior (3). Thus, pre-T1w and post-T1w are visually compared to discover extravasation of GBCA into the perivascular space (4).

This procedure has two downsides:

(I) GBCAs are generally deemed safe, but up to 2.4% of patients suffer from mild adverse reactions and a lower rate from severe complications (5). Other concerns are possible deposition of GBCAs in the brain (6), and their presence in wastewater (7). In addition, the usage of GBCAs increases scan time and monetary cost.

(II) Usage of weighted images as opposed to quantitative diagnostic methods. However, the latter are crucial for distinguishing different post-treatment conditions and understanding subtle tumor changes (4,8,9). One example is the detection of tumor infiltration into the peritumoral edema, which is difficult to assess by visual inspection of weighted images (10).

Parametric maps, specifically, longitudinal relaxation time (T1), transversal relaxation time (T2), and proton density (PD), represent intrinsic tissue properties. They are more robust than weighted images against scanner imperfections (11) and facilitate quantification of subtle changes within the tissues (4,8,12,13). Hence, parametric maps (hereafter referred to as maps) could be key for predicting BBB damage without GBCAs, as shown by Nunez-Gonzalez *et al.* (13). This study differentiates between healthy and abnormal tissue (ABN) and, particularly, tissue with and without T1w-enhancement (T1e) using only pre-contrast maps (13). However, maps are not commonly acquired in clinical practice due to their lengthy acquisitions. Nowadays, fast multiparametric mapping techniques, such as magnetic resonance (MR) fingerprinting (14) or magnetic resonance image compilation (MAGiC) (15), have taken the stage. These techniques, albeit faster, still have time limitations; moreover, these sequences are scarcely available worldwide.

Deep learning (DL) can boost parametric mapping by using only conventional and, therefore, widely available weighted images. With DL, maps can be computed in both pre-existing databases and in newly acquired datasets without additional scan time. Although various methods that estimate parametric maps from traditional relaxometry sequences with DL have been proposed (16,17), to the best of our knowledge only two previous works compute maps from conventional weighted images with DL (18,19). Moya-Sáez *et al.* (18) propose T1, T2, and PD mapping from a pre-T1w and a T2w; the method is validated on a small proprietary dataset of healthy volunteers. Qiu *et al.* (19) propose T1 and T2 mapping from three weighted images; the validation is performed with controls and multiple sclerosis patients. In both cases the maps were compared with those obtained with traditional sequences, but their diagnostic value was not clinically validated. Thus, the utility of synthesized maps for diagnostic purposes has not been so far analysed, and specifically, none of these works focused on T1e prediction in oncological patients. On the other hand, the synthesis of post-T1w images from pre-contrast weighted images has been recently proposed (20,21), but these approaches do not compute parametric maps, which impedes a quantitative assessment of the T1e.

In this pilot study, we perform a methodological feasibility analysis. It consists of predicting post-contrast T1e in glioma patients from pre-contrast conventional weighted images using DL-synthesized maps. Thus, prediction of post-contrast T1e without the usage of GBCAs alleviates the previously mentioned downside. In addition, computation of the maps with DL from conventional weighted images avoids the need of specifically targeted sequences and paves the way for retrospective studies. We present this article in accordance with the TRIPOD+AI reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-24-721/rc).

The rest of the manuscript is organized as follows: in Section Methods we present the methodology for predicting post-contrast T1e from pre-contrast conventional weighted images through the DL computation of parametric maps. Thus, we first introduce the datasets employed in this study and the pre-processing steps. Next, we introduce different approaches for T1e assessment with and without GBCAs. We present the DL approach for the synthesis of the parametric maps and the data splitting for training, early-stopping validation, and test. In the last subsection we explain the experiments carried out for the validation of the proposed approach. In Section Results we show the results whereas in Section Discussion we discuss their implications. Finally, the main conclusions of this study are established in Section Conclusions.

Methods

Figure 1 shows an overview of the proposed DL method and the experiments. Details about the datasets and number of selected participants are also included.

Datasets

In this retrospective study, we employ three different datasets (see *Figure 1A*); one of them is used for both training and testing, and the other two are solely used for testing. Details about their acquisition parameters are shown in Table S1.

Glioma patients (GLIOMA)

This dataset consists of 15 patients (mean age, 39.33±10.40 years; 6 females, 9 males) with different grades of gliomas, who were scanned with the approval of the Ethics Committee of the Erasmus MC, Rotterdam, The Netherlands (No. MEC-2018-070) and after signing a written informed consent form. The data acquisition was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The patients had undergone tumor resection before the acquisition and were scanned between 2018 and 2020 with a 3T Sigma Premier (General Electric, Waukesha, WI, USA). The image modalities included are pre-T1w, T2w, T2w-FLAIR, post-T1w, and, also, pre-contrast MAGiC[1] for T1, T2, and PD parametric mapping. In one patient the T2w-FLAIR was not acquired due to a deviation from the protocol.

Healthy volunteers relaxometry (RMaps)

Five healthy volunteers (mean age, 34.4±12.18 years; 2 females, 3 males) scanned with the approval of the Ethics Committee CEIm ÁREA DE SALUD VALLADOLID ESTE, Valladolid, Spain (No. PI-20-1885) and after signing a written informed consent form were included in this dataset. The dataset was collected in 2021 with a 3T Achieva (Philips, Best, The Netherlands). The protocol included relaxometry sequences for the estimation of T1, T2, and PD maps, and also pre-T1w and T2w images. The T1 map was estimated from a variable flip angle (VFA) sequence with 11 different flip angles using NOVIFAST (12), and the T2 map was obtained from a multi-echo sequence with six different TEs using a least squares fit.

The University of Pennsylvania glioblastoma (UPenn-GBM)

UPenn-GBM is a public dataset of 611 patients diagnosed with *de novo* glioblastoma. It includes routine clinical pre-operative MRI acquisitions from 2006 to 2018. For 41 of these patients a follow-up acquisition prior to a second resection is also available. Details about this dataset can be found in Bakas *et al.* (22). We selected 493 patients (mean age, 63.15±12.23 years; 197 females, 296 males) with homogeneous acquisition parameters. The acquisition of all the selected patients was performed with a 3T Magnetom Trio (Siemens Healthcare, Erlangen, Germany). This dataset includes four structural weighted images (pre-T1w, T2w, T2w-FLAIR, and post-T1w) resampled to a 1 mm^3 isotropic resolution and skull-stripped.

Image processing and ROI definition

All of the weighted images were reoriented to match the orientation of the standard MNI152 (23) with FSL (24). After that, images were skull-stripped using HD-BET (25) followed by a linear registration to the pre-T1w using FLIRT (24). Finally, each weighted image was normalized

---

[1] Commercial implementation of the QRAPMASTER technique for quantitative imaging.

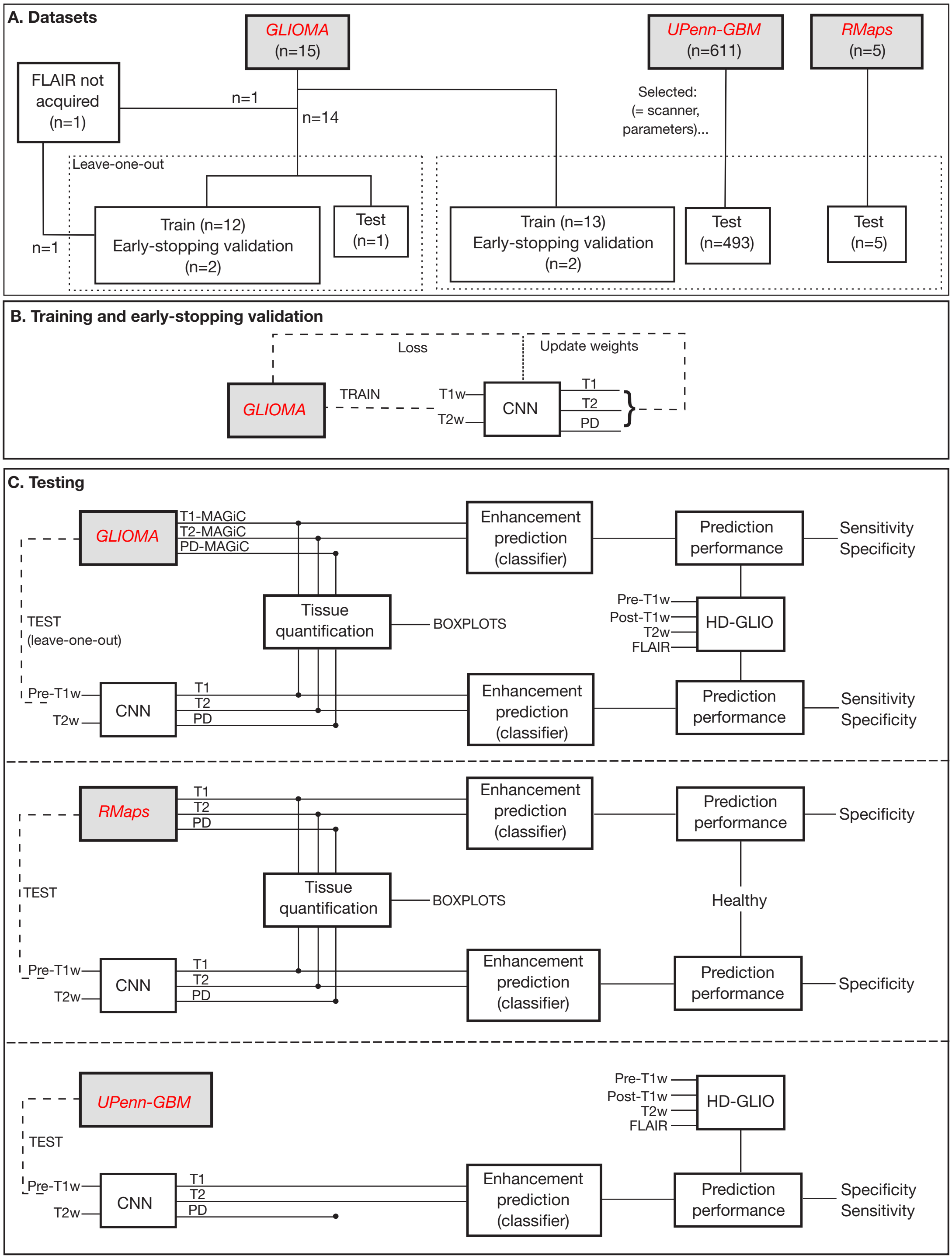

Figure 1 Flow diagram. (A) Participants selection of each dataset, namely GLIOMA, UPenn-GBM, and RMaps, and training/early-stopping validation/test sets splitting. (B) Pipeline of the proposed approach for training the DL method with the GLIOMA dataset. (C) Testing and experimentation employed for each dataset depends on the data available in each of them. CNN, convolutional neural network; T1, longitudinal relaxation time; T2, transversal relaxation time; PD, proton density; T1w, T1-weighted; T2w, T2-weigheted; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery; post-T1w, post-contrast T1-weighted; DL, deep learning; MAGiC, magnetic resonance image compilation.

with its average intensity (excluding background). FAST (24) was used to segment the white matter (WM) and grey matter (GM) tissues. In addition, different regions of interest (ROIs) were defined for the subsequent analysis: T1e, non-enhancing T2w/T2w-FLAIR signal hyperintensity (T2h), ABN, which corresponds to the union of T1e and T2h, and normal white matter (nWM), which corresponds to the WM segmentation not included in ABN.

T1e assessment

Different methods have been proposed for automatic classification of voxels with and without T1e. We will use the two we now describe.

Classification method 1: HD-GLIO

HD-GLIO is a DL segmentation tool (26-28) which uses three pre-contrast (pre-T1w, T2w, T2w-FLAIR) and one post-contrast (post-T1w) weighted images as inputs. It segments voxels with T1e, and voxels with T2h. We should highlight that this tool was trained with annotations performed by experts. In this work, GLIOMA and UPenn-GBM datasets were input to HD-GLIO for obtaining ground-truth segmentations.

Classification method 2: voxel-wise statistical prediction from pre-contrast maps

Nunez-Gonzalez *et al.* (13) recently proposed a method for voxel-wise classification of normal/ABN as well as tissue with/without T1e, based only on pre-contrast (i.e., without GBCAs) T1 and T2 maps obtained with MAGiC. Further details on this method can be found in Appendix 1. Following this methodology, in this work we considered two classifications: classification-I (C-I) ABN *vs.* nWM, and classification-II (C-II) T1e *vs.* the union of nWM and T2h. Thus, the T2h can be obtained as the difference between ABN and T1e. For each classification problem, we selected the metric that showed the best performance in Nunez-Gonzalez *et al.* (13) and its optimal operating point, i.e., normlog with threshold 8.44 for C-I, and normT1T2 with threshold 1,344 ms for C-II.

DL computation of synthesized maps

The computation of synthesized maps is performed with a convolutional neural network (CNN) that extracts the quantitative T1, T2, and PD information embedded in the weighted images (see *Figure 1B*). The CNN was configured with two encoders—one per input (pre-T1w and T2w), and three decoders—one per desired map. The pipeline has three steps: (I) each weighted image is input to its encoder; (II) the outputs of the encoders are fused with a pixel-wise max function into a shared representation; (III) the shared representation is input to each decoder. The network processes the volumes slice-wise.

The three encoders share the same architecture, which is inspired by the UNet (29) to exploit information at larger spatial scales. Skip connections were used between the down- and upsampling layers of the encoders to avoid the loss of details induced in the downsampling path. The three decoders also share the same architecture; in this case a fully-convolutional network is employed. More details about the architecture can be found in (18,30) and in the Figure S1.

Both encoders and decoders were trained simultaneously. Training was supervised with the loss function:

$$L = \|T1_{DL} - T1_{MAGiC}\|_{l1} + \|T2_{DL} - T2_{MAGiC}\|_{l1} + \|PD_{DL} - PD_{MAGiC}\|_{l1} \qquad [1]$$

where the subindices *DL* and *MAGiC* refer to the DL-synthesized and MAGiC maps, respectively. The *l1*-norm was chosen for robustness against misregistration of the input weighted images (more details in Appendix 1). The loss function was minimized using Adam (31) with a learning rate of $1e^{-4}$. The batchsize was empirically set to 4 slices. The number of epochs was determined in execution time via early-stopping (10 epochs with a loss reduction less than 0.001 in the early-stopping validation set). Pre-training with synthetic data is performed following the steps detailed in Moya-Sáez *et al.* (18). We implemented our model with Tensorflow v.2.4.0.

Training, early-stopping validation and test sets

First, the DL method was trained following a leave-one-out scheme with GLIOMA; hence, one patient was left aside for testing and the remaining patients were randomly split into training (12 patients) and early-stopping validation (two patients). The patient without T2w-FLAIR was not included in any test set because HD-GLIO cannot be executed on it, but it was additionally included in the training/early-stopping validation sets as shown in *Figure 1A*. Thus, a total of 14 data splits were performed. Patients may exhibit T1e or not; we ensured the presence of a patient of each class in the early-stopping validation set. For the sake of completeness, Table S2 includes the specific patient indexes included in each split.

Second, to assess the generalization capability of the CNN, we performed a new training with the complete GLIOMA dataset, i.e., no patient was left aside for testing. Specifically, we used one random split of GLIOMA in training (13 patients) and early-stopping validation (two patients) sets. Testing was carried out on two separate datasets (UPenn-GBM and RMaps).

Validation of the proposed approach

Two kinds of experiments have been conducted on the test sets for the validation of the proposed approach. These experiments are labelled in *Figure 1C* as tissue quantification and prediction performance. The former is intended to evaluate similarity between the synthesized and the acquired maps. The latter evaluates the capability of synthesized maps for predicting ABN and T1e tissues. We stress that this validation does not correspond with the early-stopping validation described before, which is performed during network training on the early-stopping validation set.

On the one hand, validation of the proposed approach was performed on GLIOMA using the test patient for each trained model in the leave-one-out process. As for the tissue quantification, MAGiC and synthesized maps were visually compared, and the percentage error was voxel-wise computed as the difference between them normalized by the MAGiC map. The WM and GM segmentations for each test patient were applied to both types of maps. Then, boxplots of the voxel-wise percentage error between the synthesized and MAGiC maps for the WM and GM tissues were constructed. For the prediction performance experiment, statistical predictions were carried out with both MAGiC and our synthesized maps for GLIOMA, and their sensitivities and specificities were compared. These parameters were computed for each classification problem (C-I and C-II) performed with the metrics and thresholds reported in section "Classification method 2: voxel-wise statistical prediction from pre-contrast maps". The ground-truth segmentation for computing sensitivity and specificity was obtained with HD-GLIO from the weighted images (see *Figure 1C*). Statistical differences were calculated by means of a significance test.

On the other hand, the dashed horizontal lines in *Figure 1C* depict the tests carried out on RMaps and UPenn-GBM datasets to assess the generalization capability. The tissue quantification experiment was also conducted on RMaps. In this case, the values of the parametric maps within the WM and GM segmentations were compared between our synthesized maps and the maps in RMaps by means of the voxel-wise difference between them normalized by their voxel-wise mean. As for the prediction performance, it was conducted on both datasets. For RMaps only specificity on the predictions was calculated since the dataset consists of healthy volunteers. Differences are also tested by a significance statistical test. For UPenn-GBM, predictions were performed on the synthesized maps and, afterwards, sensitivity and specificity were computed using HD-GLIO as the reference segmentation.

For the three databases, normality of the samples of sensitivity and specificity are tested with a Shapiro test. For normal data distributions, values are reported as means and standard deviation (SD). Otherwise, values are reported as median and interquartile range (IQR). As for statistical tests, either the paired *t*-test or Wilcoxon signed rank test are used according to the normality of the samples. P values are reported for a unilateral test (i.e., the alternative hypothesis is "the value obtained for our maps are greater than for the other method maps").

Results

DL-synthesized maps in GLIOMA dataset

Figure 2 shows, for a representative test patient, both the MAGIC maps and the synthesized maps. The voxel-wise percentage error between each pair of maps is also shown. An extended version of *Figure 2* with all the GLIOMA patients was provided in Figure S2. A comparison of tissue values of the WM and GM between the synthesized and MAGiC maps is shown in *Figure 3*.

T1e prediction in GLIOMA

Table 1 shows the sensitivity and specificity obtained for each classification (i.e., C-I and C-II defined in section "Classification method 2: voxel-wise statistical prediction from pre-contrast maps") for both MAGiC and synthesized maps. As can be seen in that table, synthesized maps obtained similar performance compared to MAGiC for discriminating normal from ABN (C-I) with sensitivities: 88.37% (IQR, 17.30%) *vs.* 89.35% (IQR, 8.93%) and slightly better discriminating tissue with and without T1e (C-II) with sensitivities: 93.26% (IQR, 14.05%) *vs.* 87.29% (IQR, 16.62%) and P=0.0015. Also for both maps, the segmented T1e and T2h regions obtained through the

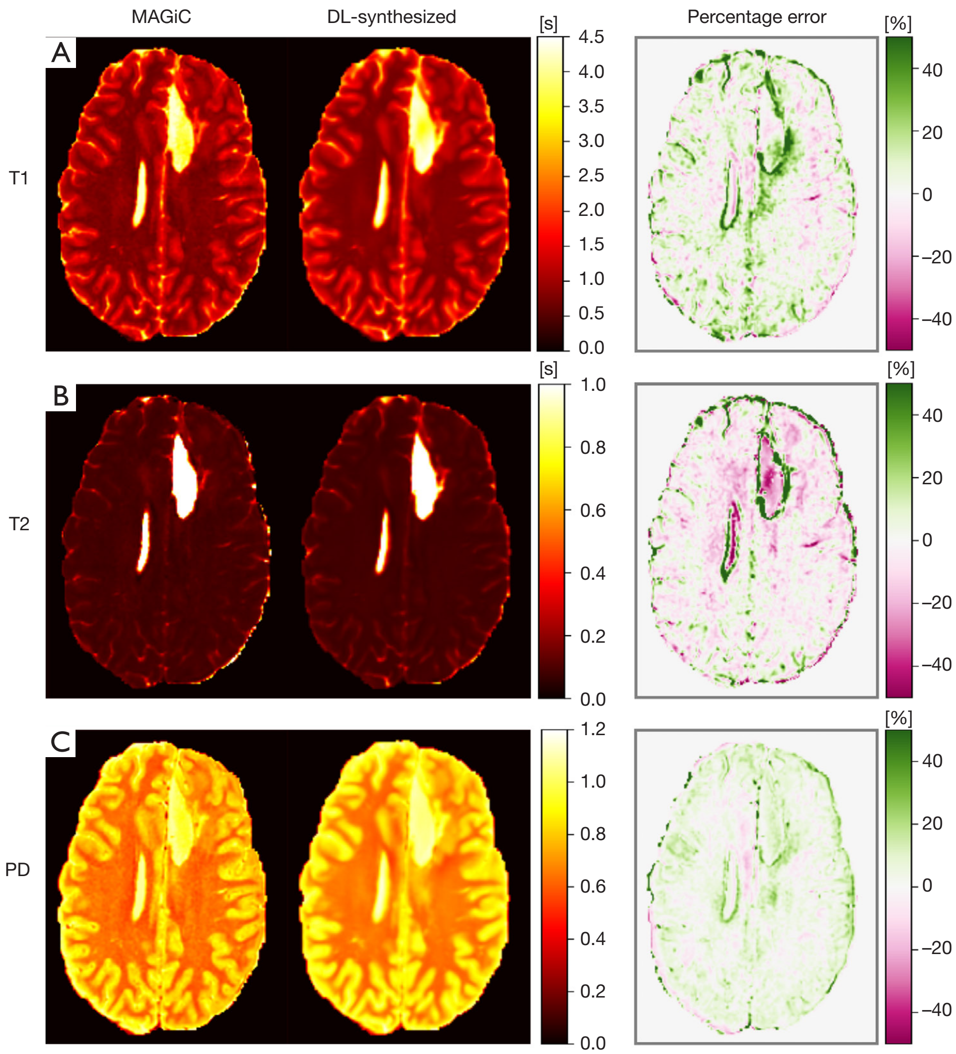

Figure 2 A representative axial slice of MAGiC and corresponding synthesized parametric maps for a test patient of GLIOMA dataset, who presents T1w-enhanced tissues. The voxel-wise percentage error is also represented for each pair. (A) T1 map; (B) T2 map; and (C) PD map. T1, longitudinal relaxation time; T2, transversal relaxation time; PD, proton density; MAGiC, magnetic resonance image compilation; DL, deep learning.

statistical prediction from both MAGiC and synthesized maps are shown in *Figure 4* for a representative test patient. The segmentations obtained with HD-GLIO from weighted images are also shown as reference. An extended version of *Figure 4* with all the GLIOMA patients was provided in Figure S3.

Results on RMaps and UPenn-GBM

Boxplots of the WM and GM percentage error between the synthesized maps and those in RMaps are shown in *Figure 5*. On the other hand, the voxel-wise statistical prediction applied to the synthesized maps reported mean specificities of 95.83% (SD: 0.94%) and 98.83% (SD: 0.16%) for C-I and C-II, respectively, while for the relaxometry maps from RMaps the specificity lowered to 82.08% (SD: 5.93%) for C-I and 83.36% (SD: 6.03%) for C-II. Significant differences were found in both cases (P value <0.01).

Figure 6 shows a comparison between the segmentation obtained with the voxel-wise statistical prediction from DL-synthesized maps and the reference HD-GLIO segmentation for UPenn-GBM. Sensitivities and specificities for the voxel-wise prediction with the

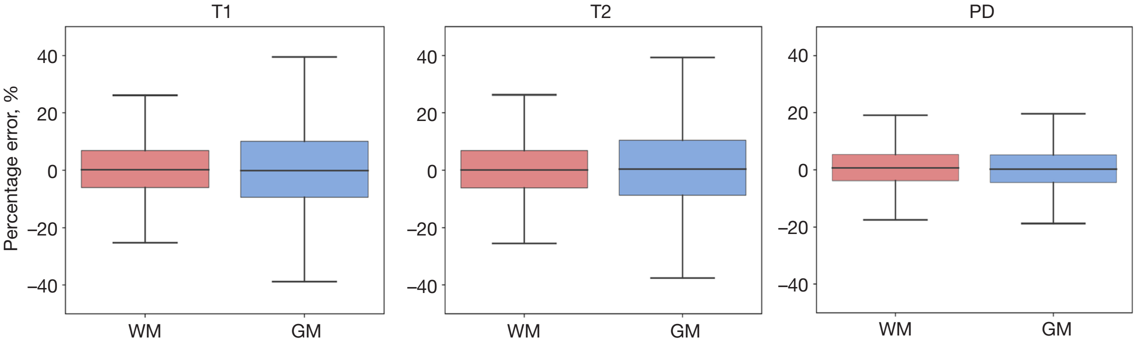

Figure 3 Boxplots of the voxel-wise percentage error between the synthesized and the MAGiC parametric maps within the WM and GM segmentations for the test patients of the GLIOMA dataset. T1, longitudinal relaxation time; T2, transversal relaxation time; PD, proton density; MAGiC, magnetic resonance image compilation; WM, white matter; GM, grey matter.

Table 1 Comparison of the statistical prediction performance with MAGiC and synthesized maps for test patients of GLIOMA dataset

Metric	C-I: ABN *vs.* nWM	C-II: T1e *vs.* non-T1e
Sensitivity, %		
MAGiC	89.35 (8.93)	87.29 (16.62)
Synthesized	88.37 (17.30)	93.26 (14.05)
P value	0.4039	**0.0015**
Specificity, %		
MAGiC	93.61 (2.14)	94.95 (2.88)
Synthesized	95.21 (1.13)	95.59 (1.81)
P value	**0.0011**	**0.0054**

Median across patients and interquartile range for each classification problem (C-I and C-II) are reported. Bold text represents significant differences (P value <0.05) between synthesized and MAGiC maps. ABN, abnormal tissue; nWM, normal white matter; T1e, T1w-enhancement; MAGiC, magnetic resonance image compilation.

synthesized maps are shown in *Table 2*. The metrics reported are sensitivities of 91.23% (IQR, 10.14%) and 81.04% (IQR, 24.21%) and specificities of 90.24% (IQR, 3.55%) and 91.49% (IQR, 6.14%) for C-I and C-II, respectively. These values are comparable to those reported in *Table 1* for GLIOMA.

Discussion

In this pilot study, we propose a method that enables the prediction of brain tumor T1e from pre-contrast conventional weighted images through the computation of synthesized maps. We validate the ability of these maps to predict ABN and T1e tissues without GBCAs and, also, their tissue quantification accuracy. Generalization capabilities of the proposed method were also tested with multi-site, multi-vendor acquisitions.

Results have shown that the DL-based parametric mapping from only two weighted images show tissue values similar to those obtained with relaxometry sequences and MAGiC. In addition, the performance of these synthesized maps to discriminate normal/ABN is comparable to the performance obtained with MAGiC maps and slightly better for tissue with or without T1e, both on the GLIOMA dataset. These results could be replicated by testing our method with RMaps and UPenn-GBM, which were two datasets unseen during the training process. The three datasets were collected in different centers and with 3T scanners of multiple vendors. Neither inhomogeneity nor motion correction techniques were applied to the estimation of the maps of the RMaps dataset, which could affect their accuracy. This could have an impact on the better performance of the synthesized maps in the statistical predictions. In contrast, MAGiC maps were obtained with a commercial product which incorporates B1 inhomogeneities into the model for T1, T2, and PD mapping. In addition, its accuracy has been validated in various studies (32,33).

Interestingly, the statistical prediction shows a slight overestimation of the T1e region compared to HD-GLIO, but there is also some coincidence with the follow-up images shown in *Figure 6*. This region presents T1 and T2 values higher than those of the nWM, resulting in the region being classified as T1e. Nevertheless, we hypothesize that these higher values could be explained by an altered interstitial fluid mobility and increased water content in the perivascular space (34), even when the BBB

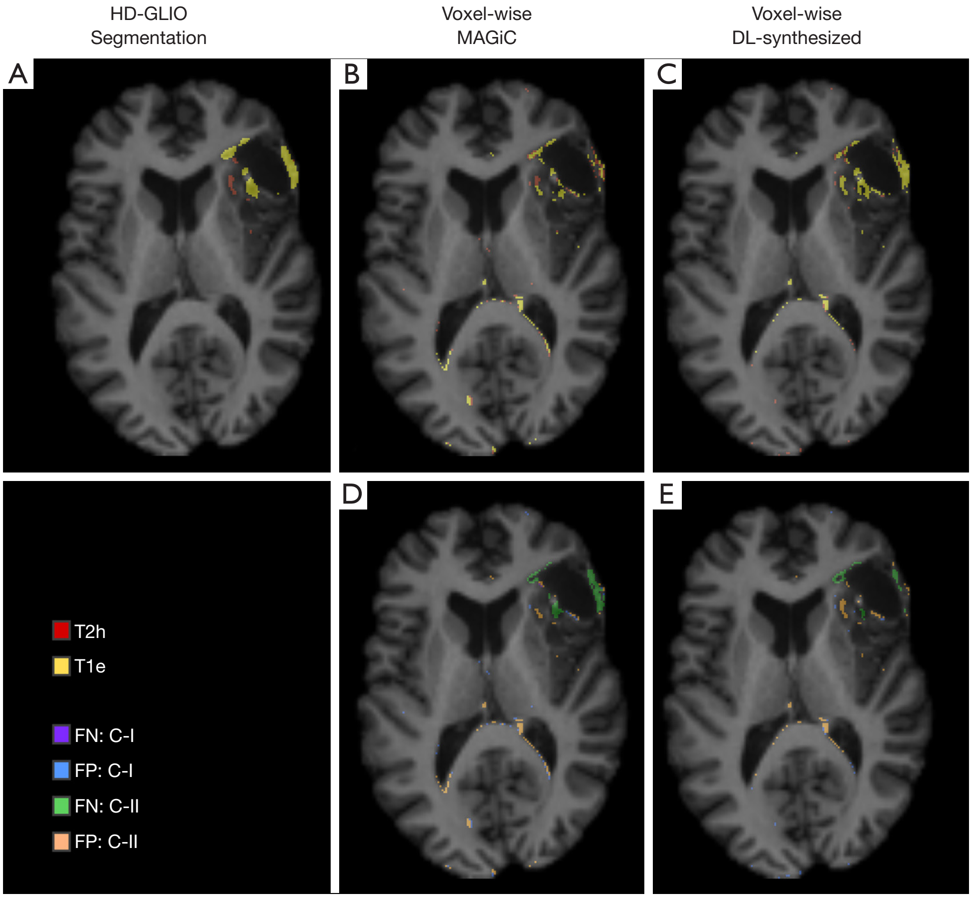

Figure 4 A representative axial slice of the T1e and T2h segmentations overlaid on the pre-T1w images for a representative test patient of the GLIOMA dataset. (A) Ground-truth segmentation obtained with HD-GLIO from the four weighted images (pre-contrast T1w, T2w, T2w-FLAIR, and post-T1w). (B) Segmentations obtained through the statistical predictions from MAGiC parametric maps. (C) Segmentations obtained through the statistical predictions performed from synthesized maps. (D) Errors of B compared with A. (E) Errors of C compared to A. FN, false negative; FP, false positive; T1w, T1-weighted; T2w, T2-weigheted; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery; post-T1w, post-contrast T1-weighted; MAGiC, magnetic resonance image compilation; T1e, T1w-enhancement; T2h, T2w/T2w-FLAIR signal hyperintensity; DL, deep learning.

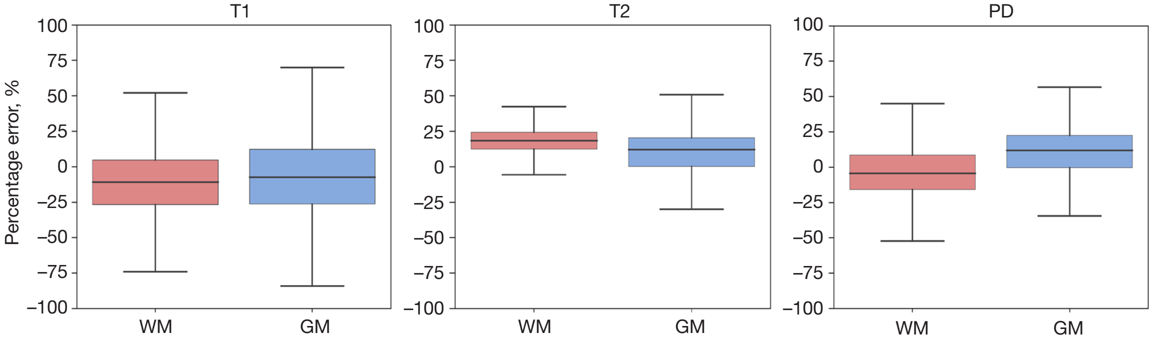

Figure 5 Boxplots of the voxel-wise percentage error between the synthesized and the relaxometry maps of RMaps dataset within the WM and GM mask. T1, longitudinal relaxation time; T2, transversal relaxation time; PD, proton density; WM, white matter; GM, grey matter.

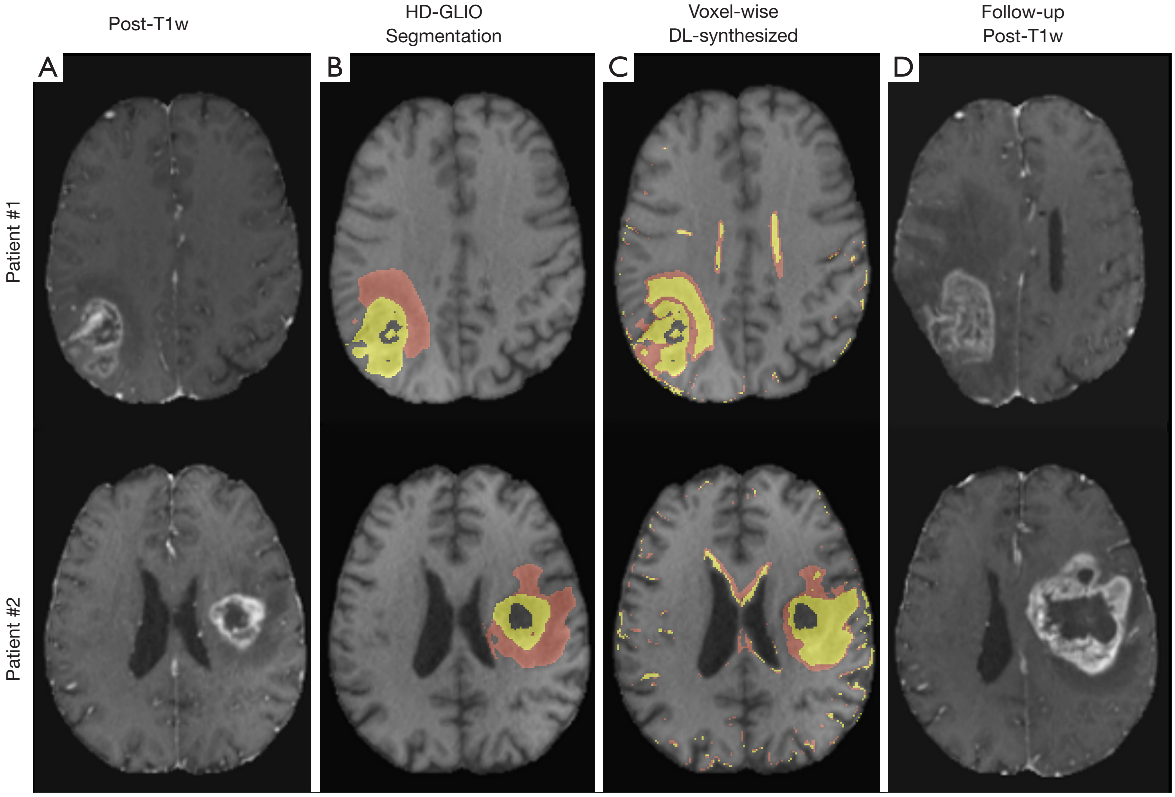

Figure 6 A representative axial slice of the T1e and T2h segmentations overlaid on the pre-T1w images for different patients of UPenn-GBM dataset. (A) Post-T1w image. (B) Reference segmentation obtained with HD-GLIO from the four weighted images (pre-contrast T1w, T2w, T2w-FLAIR, and post-T1w). (C) Segmentations obtained through the statistical predictions performed from synthesized maps. (D) Post-T1w image from the follow-up acquisition. Red label represents T2h, whereas the yellow label represents T1e. T1w, T1-weighted; T2w, T2-weigheted; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery; post-T1w, post-contrast T1-weighted; T1e, T1w-enhancement; T2h, T2w/T2w-FLAIR signal hyperintensity; DL, deep learning.

Table 2 Statistical prediction performance of synthesized maps for patients of UPenn-GBM dataset in terms of sensitivity and specificity

Metric	C-I: ABN *vs.* nWM	C-II: T1e *vs.* non T1e
Sensitivity, %	91.23 (10.14)	81.04 (24.21)
Specificity, %	90.24 (3.55)	91.49 (6.14)

Median across patients and interquartile range for each classification problem (C-I and C-II) are reported. ABN, abnormal tissue; nWM, normal white matter; T1e, T1w-enhancement.

is not completely disrupted for GBCA to pass through. Thus, this overestimated T1e region could represent a BBB vulnerability and, therefore, it might be an appropriate follow-up biomarker. This hypothesis agrees both with recently reported findings (8,35) and, apparently, with our results on the UPenn-GBM follow-up images (see *Figure 6*). However, further research is necessary to confirm this hypothesis.

The prediction of the T1e from pre-contrast maps could bypass the injection of GBCAs, avoiding their related issues, as shown in Nunez-Gonzalez *et al.* (13). However, the acquisition of maps makes protocols considerably longer. MAGiC does not fully address this issue since its acquisition is still not less than 4 min for a resolution of 1 mm × 1 mm × 3 mm. Thus, computation of maps from conventional weighted images with DL would favor the spread of this clinical application and foster the usage of maps in the clinical practice. Moreover, this approach proves valuable for retrospective analysis, enabling the computation of maps for pre-existing and/or multicentre datasets, thereby empowering longitudinal or population studies.

If it is feasible to derive maps from the weighted images and subsequently predict T1e from these maps, it follows that predicting T1e directly from the weighted images should also be viable. However, we choose to use the maps as an intermediate step for several reasons: (I) gaining resilience to variations in scanners and acquisition parameters by extracting the absolute scale maps; (II) offering information closer to the pathophysiology; and (III) providing DL-based evidence of the findings reported in previous studies (8,13,35,36) which have shown that T1e tissues present higher pre-contrast T1 and T2 values compared with nWM.

This work has several limitations. Acquisition time limitations lead to a mismatch in spatial resolution between the weighted images and MAGiC maps of the GLIOMA dataset. Consequently, the co-registration of all modalities is needed, which might induce misregistration artifacts and partial volume effects. This can explain why most of the differences between the synthesized and MAGiC maps are located at the tissue edges (see *Figure 2*). Additionally, the network might be inducing some blurring in the synthesized maps, which could also explain these tissue edge differences. Moreover, the pre-T1w and T2w input images have varying acquisition parameters across datasets due to different institutional protocols. Although this study has shown comparable testing results on diverse datasets, these input parameter variations could potentially introduce errors in the maps.

To mitigate the impact of variation in acquisition parameters, future work aims to incorporate them as additional inputs to the network. This approach would enable training with more extensive and diverse datasets, potentially enhancing the model generalization capabilities. Furthermore, the extension of the approach to other clinical applications for diagnosing different pathologies is also planned (37).

Conclusions

In conclusion, in this study we showed a proof-of-concept of predicting T1e from pre-contrast conventional weighted images through the DL synthesis of parametric maps. The results suggest that these DL-synthesized maps might eventually have the potential to replace GBCAs for tumor T1e prediction, without compromising performance and with a reduced acquisition time.

Acknowledgments

Funding: This work was supported by the Fundación Científica Asociación Española Contra el Cáncer (FC AECC), the Ministerio de Ciencia e Innovación of Spain (Nos. PID2020-115339RB-I00, TED2021-130090B-I00, and PID2021-124407NB-I00), the General Electric Healthcare (grant: “B-GEHC-5. MR Physiological Signature”), and ESAOTE ESPAÑA (No. 18I QBM). In addition, the project leading to these results has received funding from “la Caixa” Foundation and FCT, I.P. under the project code (No. LCF/PR/HR22/00533).

Footnote

Reporting Checklist: The authors have completed the TRIPOD+AI reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-24-721/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-24-721/coif). J.A.H.T. has received funding for research from GE HealthCare. C.A.L. has received funding for research from ESAOTE ESPAÑA. E.M.S. has received a predoctoral fellowship from Fundación Científica Asociación Española Contra el Cáncer. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of the Erasmus MC, Rotterdam, The Netherlands (No. MEC-2018-070), and Ethics Committee CEIm ÁREA DE SALUD VALLADOLID ESTE, Valladolid, Spain (No. PI-20-1885). The subjects were scanned after signing a written informed consent form.

References

1. Wen PY, Kesari S. Malignant gliomas in adults. N Engl J Med 2008;359:492-507.
2. Thust SC, Heiland S, Falini A, Jäger HR, Waldman AD, Sundgren PC, Godi C, Katsaros VK, Ramos A, Bargallo N, Vernooij MW, Yousry T, Bendszus M, Smits M. Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice. Eur Radiol 2018;28:3306-17.
3. Warntjes M, Blystad I, Tisell A, Larsson EM. Synthesizing a Contrast-Enhancement Map in Patients with High-Grade Gliomas Based on a Postcontrast MR Imaging Quantification Only. AJNR Am J Neuroradiol 2018;39:2194-9.
4. Hattingen E, Müller A, Jurcoane A, Mädler B, Ditter P, Schild H, Herrlinger U, Glas M, Kebir S. Value of quantitative magnetic resonance imaging T1-relaxometry in predicting contrast-enhancement in glioblastoma patients. Oncotarget 2017;8:53542-51.
5. Forghani R. Adverse Effects of Gadolinium-Based Contrast Agents: Changes in Practice Patterns. Top Magn Reson Imaging 2016;25:163-9.
6. Yao X, Hu J, Wang G, Lin X, Sun J, Dong G, Kang J, Feng W, Xie B, Huang Y, Tian X, Chen E. Deposition of Gadolinium in the Central and Peripheral Nervous Systems and Its Effects on Sensory, Cognitive, and Athletic Implications after Multiple Injections of Gadolinium-Based Contrast Agents in Rats. AJNR Am J Neuroradiol 2024;45:1153-61.
7. Inoue K, Fukushi M, Furukawa A, Sahoo SK, Veerasamy N, Ichimura K, Kasahara S, Ichihara M, Tsukada M, Torii M, Mizoguchi M, Taguchi Y, Nakazawa S. Impact on gadolinium anomaly in river waters in Tokyo related to the increased number of MRI devices in use. Mar Pollut Bull 2020;154:111148.
8. Lescher S, Jurcoane A, Veit A, Bähr O, Deichmann R, Hattingen E. Quantitative T1 and T2 mapping in recurrent glioblastomas under bevacizumab: earlier detection of tumor progression compared to conventional MRI. Neuroradiology 2015;57:11-20.
9. Pirkl CM, Nunez-Gonzalez L, Kofler F, Endt S, Grundl L, Golbabaee M, Gómez PA, Cencini M, Buonincontri G, Schulte RF, Smits M, Wiestler B, Menze BH, Menzel MI, Hernandez-Tamames JA. Accelerated 3D whole-brain T1, T2, and proton density mapping: feasibility for clinical glioma MR imaging. Neuroradiology 2021;63:1831-51.
10. Blystad I, Warntjes JBM, Smedby Ö, Lundberg P, Larsson EM, Tisell A. Quantitative MRI using relaxometry in malignant gliomas detects contrast enhancement in peritumoral oedema. Sci Rep 2020;10:17986.
11. Weiskopf N, Suckling J, Williams G, Correia MM, Inkster B, Tait R, Ooi C, Bullmore ET, Lutti A. Quantitative multi-parameter mapping of R1, PD(*), MT, and R2(*) at 3T: a multi-center validation. Front Neurosci 2013;7:95.
12. Badve C, Yu A, Dastmalchian S, Rogers M, Ma D, Jiang Y, Margevicius S, Pahwa S, Lu Z, Schluchter M, Sunshine J, Griswold M, Sloan A, Gulani V. MR Fingerprinting of Adult Brain Tumors: Initial Experience. AJNR Am J Neuroradiol 2017;38:492-9.
13. Nunez-Gonzalez L, van Garderen KA, Smits M, Jaspers J, Romero AM, Poot DHJ, Hernandez-Tamames JA. Pre-contrast MAGiC in treated gliomas: a pilot study of quantitative MRI. Sci Rep 2022;12:21820.
14. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, Griswold MA. Magnetic resonance fingerprinting. Nature 2013;495:187-92.
15. Warntjes JB, Leinhard OD, West J, Lundberg P. Rapid magnetic resonance quantification on the brain: Optimization for clinical usage. Magn Reson Med 2008;60:320-9.
16. Sabidussi ER, Klein S, Caan MWA, Bazrafkan S, den Dekker AJ, Sijbers J, Niessen WJ, Poot DHJ. Recurrent inference machines as inverse problem solvers for MR relaxometry. Med Image Anal 2021;74:102220.
17. Liu F, Kijowski R, El Fakhri G, Feng L. Magnetic resonance parameter mapping using model-guided self-supervised deep learning. Magn Reson Med 2021;85:3211-26.
18. Moya-Sáez E, Peña-Nogales Ó, Luis-García R, Alberola-López C. A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data. Comput Methods Programs Biomed 2021;210:106371.
19. Qiu S, Chen Y, Ma S, Fan Z, Moser FG, Maya MM, Christodoulou AG, Xie Y, Li D. Multiparametric mapping in the brain from conventional contrast-weighted images using deep learning. Magn Reson Med 2022;87:488-95.
20. Moya-Sáez E, de Luis-García R, Alberola-López C. Toward deep learning replacement of gadolinium in neuro-oncology: A review of contrast-enhanced synthetic MRI. Front Neuroimaging 2023;2:1055463.
21. Calabrese E, Rudie JD, Rauschecker AM, Villanueva-Meyer JE, Cha S. Feasibility of Simulated Postcontrast MRI of Glioblastomas and Lower-Grade Gliomas by Using Three-dimensional Fully Convolutional Neural

Networks. Radiol Artif Intell 2021;3:e200276.
22. Bakas S, Sako C, Akbari H, Bilello M, Sotiras A, Shukla G, et al. The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics. Sci Data 2022;9:453.
23. Evans AC, Janke AL, Collins DL, Baillet S. Brain templates and atlases. Neuroimage 2012;62:911-22.
24. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. Neuroimage 2012;62:782-90.
25. Schell M, Tursunova I, Fabian I, Bonekamp D, Neuberger U, Wick W, Bendszus M, Maier-Hein K, Kickingereder P. Automated brain extraction of multi-sequence MRI using artificial neural networks. In proceedings of European Congress of Radiology-ECR 2019. doi: 10.26044/ecr2019/C-2088.
26. Kickingereder P, Isensee F, Tursunova I, Petersen J, Neuberger U, Bonekamp D, et al. Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. Lancet Oncol 2019;20:728-40.
27. Isensee F, Petersen J, Kohl SA, Jäger PF, Maier-Hein KH. nnu-net: Breaking the spell on successful medical image segmentation. arXiv preprint 2019;arXiv:1904.08128,1(1-8):2.
28. HD-GLIO. Available online: https://github.com/NeuroAI-HD/HD-GLIO (Accessed: 16 July 2024).
29. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015;Part III:234-41.
30. Chartsias A, Joyce T, Giuffrida MV, Tsaftaris SA. Multimodal MR Synthesis via Modality-Invariant Latent Representation. IEEE Trans Med Imaging 2018;37:803-14.
31. Jais IKM, Ismail AR, Nisa SQ. Adam optimization algorithm for wide and deep neural network. Knowl Eng Data Sci 2019;2:41-6.
32. Hagiwara A, Hori M, Cohen-Adad J, Nakazawa M, Suzuki Y, Kasahara A, Horita M, Haruyama T, Andica C, Maekawa T, Kamagata K, Kumamaru KK, Abe O, Aoki S. Linearity, Bias, Intrascanner Repeatability, and Interscanner Reproducibility of Quantitative Multidynamic Multiecho Sequence for Rapid Simultaneous Relaxometry at 3 T: A Validation Study With a Standardized Phantom and Healthy Controls. Invest Radiol 2019;54:39-47.
33. Nunez-Gonzalez L, Kotek G, Gómez PA, Buonincontri G, Vogel M, Krestin GP, Poot DHJ, Hernandez-Tamames JA. Accuracy and repeatability of QRAPMASTER and MRF-vFA. Magn Reson Imaging 2021;83:196-207.
34. Wardlaw JM, Valdés Hernández MC, Muñoz-Maniega S. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. J Am Heart Assoc 2015;4:001140.
35. Müller A, Jurcoane A, Kebir S, Ditter P, Schrader F, Herrlinger U, Tzaridis T, Mädler B, Schild HH, Glas M, Hattingen E. Quantitative T1-mapping detects cloudy-enhancing tumor compartments predicting outcome of patients with glioblastoma. Cancer Med 2017;6:89-99.
36. Hattingen E, Jurcoane A, Daneshvar K, Pilatus U, Mittelbronn M, Steinbach JP, Bähr O. Quantitative T2 mapping of recurrent glioblastoma under bevacizumab improves monitoring for non-enhancing tumor progression and predicts overall survival. Neuro Oncol 2013;15:1395-404.
37. Deoni SC. Quantitative relaxometry of the brain. Top Magn Reson Imaging 2010;21:101-13.

Cite this article as: Moya-Sáez E, de Luis-García R, Nunez-Gonzalez L, Alberola-López C, Hernández-Tamames JA. Brain tumor enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and generative artificial intelligence. Quant Imaging Med Surg 2025;15(1):42-54. doi: 10.21037/qims-24-721

Appendix 1

Deep learning (DL) model architecture, loss function, and data splitting

The DL model architecture is shown in *Figure S1*. More details about the architecture can be found in (18,30).

The DL model was trained in a supervised manner with the loss function:

$$L = \left\| T1_{DL} - T1_{MAGiC} \right\|_{l1} + \left\| T2_{DL} - T2_{MAGiC} \right\|_{l1} + \left\| PD_{DL} - PD_{MAGiC} \right\|_{l1}$$

As can be seen, the loss function is composed of the sum of three terms, one per output parametric map [i.e., longitudinal relaxation time (T1), transversal relaxation time (T2), and proton density (PD) maps]. Each term consists of the l1 norm, also known as mean absolute error (MAE), between synthesized and the reference parametric maps. Note that in this case the reference parametric maps are the magnetic resonance image compilation (MAGiC) maps. MAE is defined as the mean of the absolute difference between both maps.

Finally, *Table S2* shows the patient indexes included in training, early-stopping validation and test sets for each split.

Voxel-wise statistical prediction of T1w-enhancement

Nunez-Gonzalez *et al.* (13) proposed a method for voxel-wise statistical prediction of normal/ABN tissue and tissue with and without T1e from only pre-contrast MAGiC parametric maps. In that study, the authors performed the voxel-wise prediction using a receiver operating characteristic curve (ROC) analysis. The authors considered three classification problems: classification-I (C-I) abnormal tissue (ABN) *vs.* normal white matter (nWM); classification-II (C-II) T1e *vs.* (nWM + T2h); and classification-III (C-III) T1e *vs.* T2h (only inside ABN). They also considered four voxel-wise metrics: (I) T1 values, (II) T2 values, (III) normT1T2 (i.e., the Euclidean norm of the T1 and T2 values), and (IV) normlog (i.e., the Euclidian norm of the logarithm of T1 and T2 values). For each classification problem the metric with the highest area under the curve (AUC) was selected. The thresholds for optimal classifications were calculated as the highest Youden's index of the ROC curve. The selected metrics and thresholds were normlog with threshold 8.44 for C-I, normT1T2 with threshold 1,344 ms for C-II, and normT1T2 with threshold 1,512 ms for C-III. The authors showed that the ROC analysis including PD did not improve the AUC in any of the cases. Consequently, they excluded the PD values from the rest of the analysis. Accordingly, we perform the classifications by applying their selected thresholds to the aforementioned metrics inside the white matter plus ABN masks to discriminate between the different regions.

Additional qualitative analysis

Figure S2 shows a representative slice of both MAGIC and synthesized maps for all patients of GLIOMA dataset. It can be noticed that most of the differences between both maps are located at the interfaces between tissues. This effect might be caused by partial volume effects and misregistration of the input weighted images. In general, PD maps present less differences than T1 and T2 maps for all patients. In patients 2 and 6 the T1 values in the resected regions are higher in the synthesized than in the MAGiC maps. We hypothesize this could be due to a poor estimation of the MAGiC maps in these two patients, given that these extremely low T1 values in the resected regions are not present across the other 12 patients.

Figure S3 shows the segmented T1e and T2h regions obtained through the statistical prediction from both MAGiC and synthesized maps for all patients of GLIOMA dataset. It is worth noting that segmented T1e and T2h regions are similar in both maps for all patients and that both present an overestimation of the T1e with respect to HD-GLIO. We hypothesize this could happen due to an altered interstitial fluid mobility and increased water content in the perivascular space.

Table S1 Details about the acquisition parameters for the different image modalities and each dataset

	Pre-T1w	T2w	T2w-FLAIR	Post-T1w	MAGiC	T1 map	T2 map	PD map
GLIOMA								
Voxel size (mm^2)	1.0×1.0	0.6×0.6	1.1×1.1	1.0×1.0	1.0×1.0			
Slice thickness (mm)	1.0	3.0	1.6	1.0	3.0			
Slice spacing (mm)	–	3.0	–	–	3.0			
# Slices	352	49	224	352	49			
FOV (mm^2)	240×240	233×233	246×246	240×240	240×240			
TE (ms)	3.3	97	89	3.3	6,114			
TR (ms)	7.9	9,837	5,000	7.9	15.7			
TI (ms)	450	–	1,588	450	11			
Flip angle (°)	12	90	90	12	90			
Scan time (min)	~5	~4	~4	~5	~5			
UPenn-GBM								
Voxel size (mm^2)	0.98×0.98	0.9×0.9	0.94×0.94	0.98×0.98				
Slice thickness (mm)	1.0	0.9	3.0	1.0				
Slice spacing (mm)	NA	NA	NA	NA				
# Slices	155	155	155	155				
FOV (mm^2)	240×240	240×240	240×240	240×240				
TE (ms)	3.1	458	140	3.1				
TR (ms)	1,760	3,200	9,420	1,760				
TI (ms)	950	–	2,500	950				
Flip angle (°)	15	120	170	15				
RMaps								
Voxel size (mm^2)	1.25×1.25	1.02×1.36				1.5×1.5	1.5×1.5	1.5×1.5
Slice thickness (mm)	1.2	3				1.5	1.5	1.5
Slice spacing (mm)	–	3				–	–	–
# Slices	170	50				150	150	150
FOV (mm^2)	240×240	260×195				240×240	240×240	240×240
TE (ms)	3	85				2	17, 46, 75, 104, 133, 162	2
TR (ms)	6.44	4,000				18	1,000	50
TI (ms)	900	–				–	–	–
Flip angle (°)	10	90				2, 3, 4, 5, 7, 9, 11, 14, 17, 19, 22	90	5
Scan time (min)	~4	2:30–4:30				~17:00	~18:00	~4:00

T1w, T1-weighted; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery; MAGiC, magnetic resonance image compilation; FOV, field of view; TR, repetition time; TE, echo time; TI, inversion time.

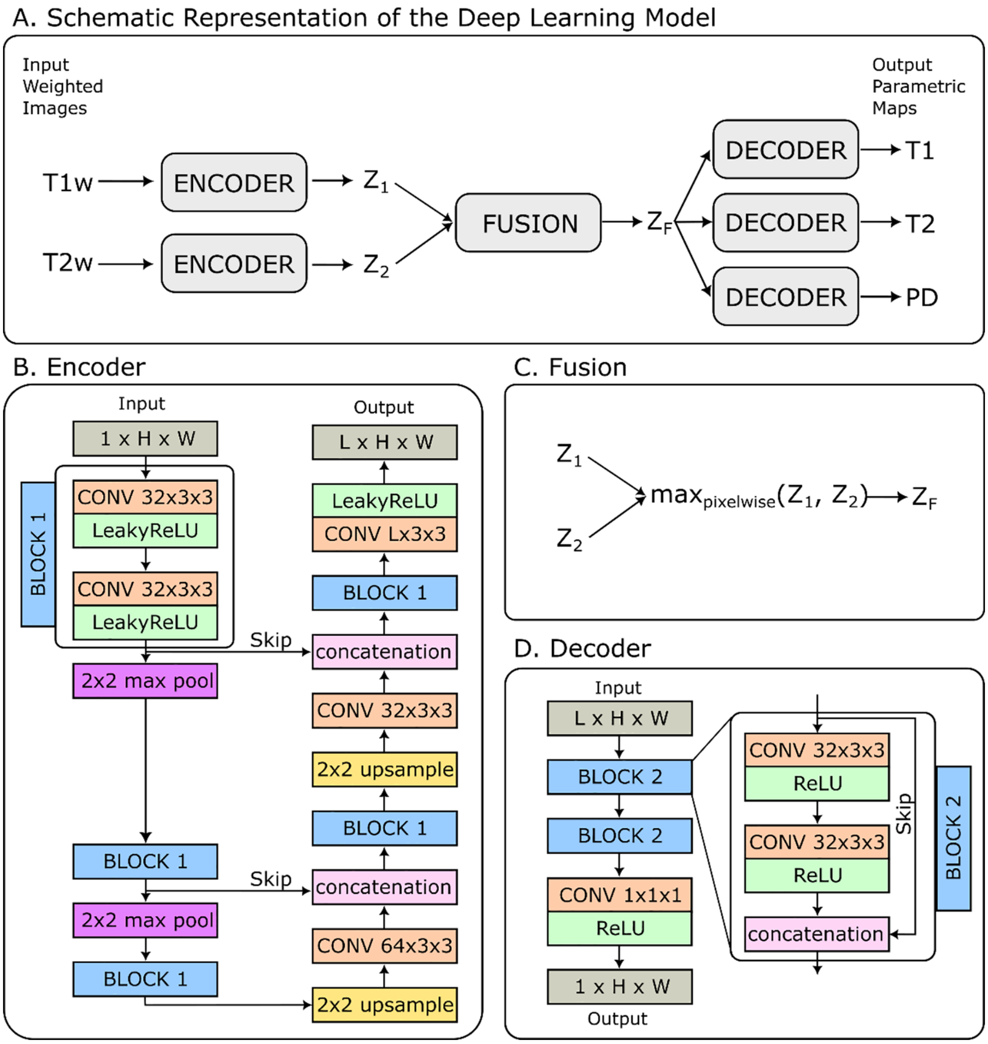

Figure S1 DL model architecture. (A) Schematic representation of the DL model. (B) Architecture of the encoder modules. (C) Fusion stage. (D) Architecture of the decoder modules. DL, deep learning.

Table S2 Indexes of patients selected for training, early-stopping validation and test of the GLIOMA dataset

Test	Early-stopping validation	Training
Test with GLIOMA (leave-one-out)		
0	1, 6	2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14
1	4, 11	0, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14
2	8, 11	0, 1, 3, 4, 5, 6, 7, 9, 10, 12, 13, 14
3	14, 11	0, 1, 2, 4, 5, 6, 7, 8, 9, 10, 12, 13
4	14, 11	0, 1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13
5	1, 11	0, 2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14
6	14, 10	0, 1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13
7	14, 10	0, 1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 13
8	5, 6	0, 1, 2, 3, 4, 7, 9, 10, 11, 12, 13, 14
9	12, 11	0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 13, 14
10	1, 6	0, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14
11	5, 6	0, 1, 2, 3, 4, 7, 8, 9, 10, 12, 13, 14
13	12, 6	0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 14
14	8, 11	0, 1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 13
Test with independent datasets		
UPenn-GBM and RMaps	1, 6	0, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14

Note that patients 6, 10, and 11 do not present T1e, and patient 12 was not included in any test set because the T2w-FLAIR was not available. T1e, T1-weighted-enhancement; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery.

https://dx.doi.org/10.21037/qims-24-721

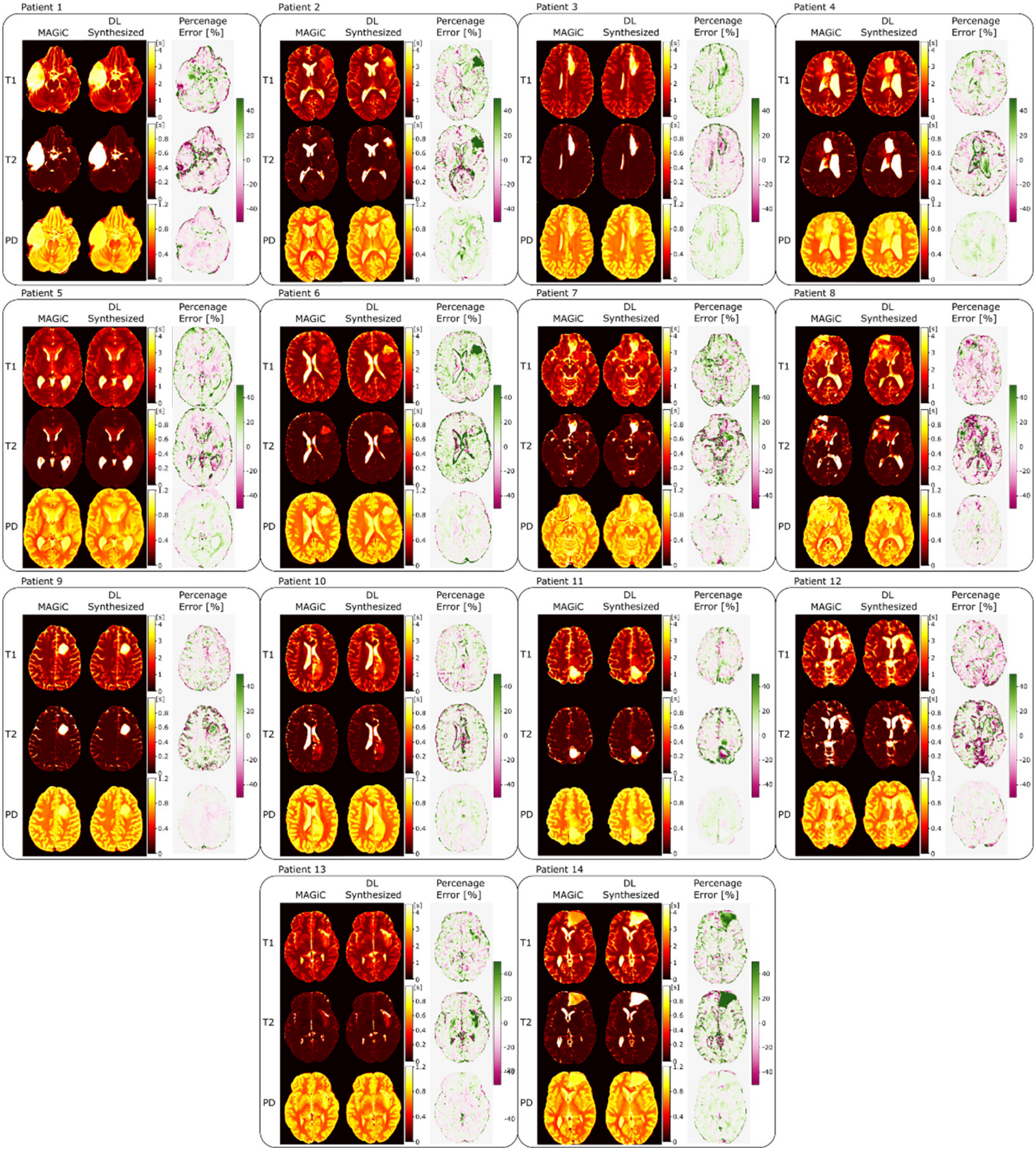

Figure S2 A representative axial slice of MAGiC and corresponding synthesized maps for the test patients of leave-one-out cross-validation with GLIOMA dataset. The voxel-wise percentage error is also represented for each pair of maps.

https://dx.doi.org/10.21037/qims-24-721

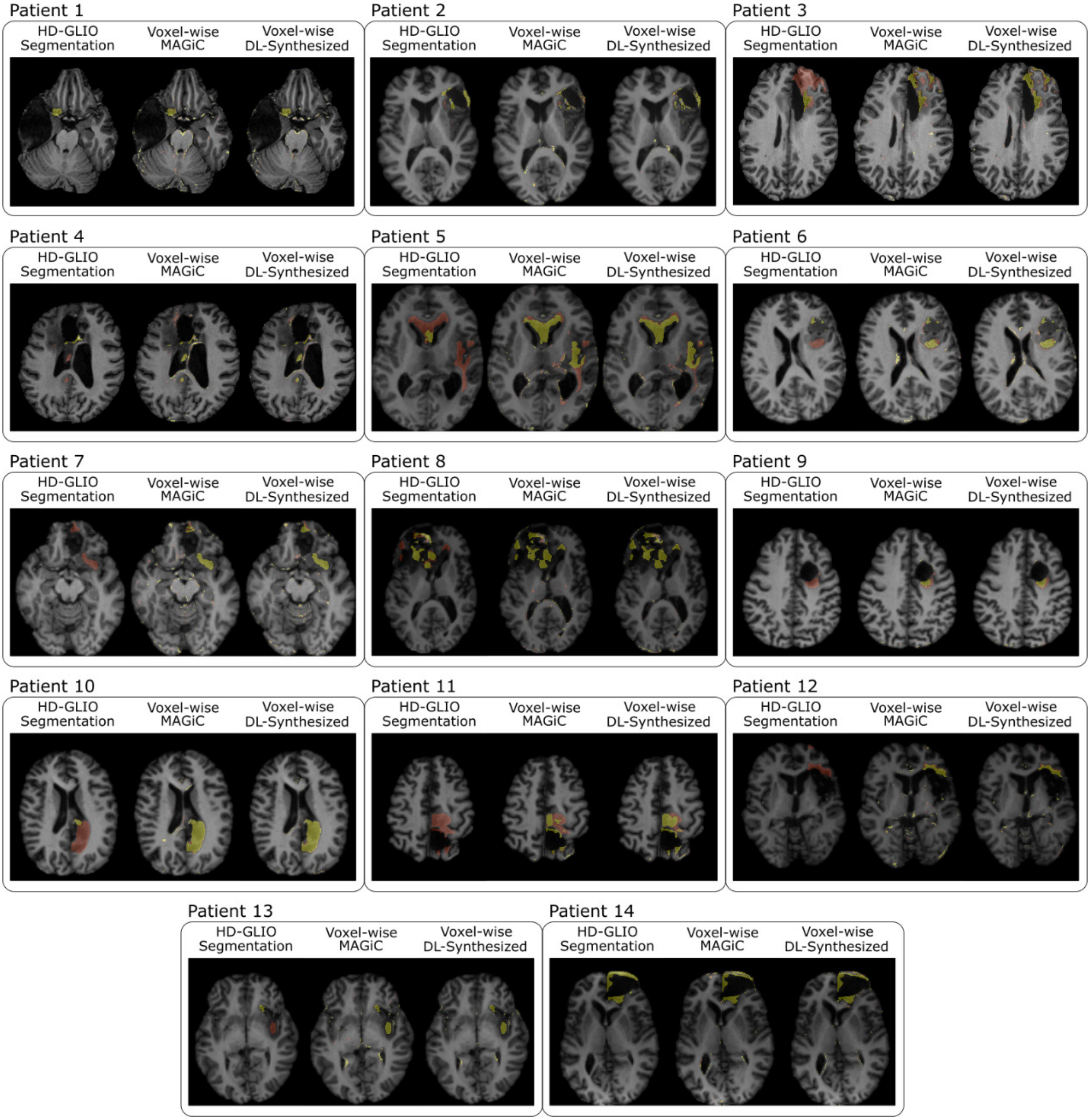

Figure S3 A representative axial slice of the T1e and T2h segmentations overlaid on the pre-T1w images for all the test patients of leave-one-out cross-validation with GLIOMA dataset. For each patient, the ground-truth segmentation obtained with HD-GLIO from the four weighted images, the segmentations obtained through the statistical predictions from MAGiC maps, and the segmentations obtained through the statistical predictions from synthesized maps are shown. MAGiC, magnetic resonance image compilation; T1e, T1w-enhancement; T2h, T2w/T2w-FLAIR signal hyperintensity; T1w, T1-weighted; FLAIR, fluid-attenuated inversion recovery.

https://dx.doi.org/10.21037/qims-24-721