Received: 17 November 2020 | Revised: 24 November 2020 | Accepted: 5 January 2021

DOI: 10.1002/nbm.4479

RESEARCH ARTICLE

NMR IN BIOMEDICINE WILEY

Deep learning magnetic resonance spectroscopy fingerprints of brain tumours using quantum mechanically synthesised data

Nikolaos Dikaios[1,2]

[1]Mathematics Research Center, Academy of Athens, Athens, Greece

[2]Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK

Correspondence
Nikolaos Dikaios, Mathematics Research Center, Academy of Athens, 11527 Athens, Greece.
Email: n.dikaios@gmail.com

Metabolic fingerprints are valuable biomarkers for diseases that are associated with metabolic disorders. 1H magnetic resonance spectroscopy (MRS) is a unique noninvasive diagnostic tool that can depict the metabolic fingerprint based solely on the proton signal of different molecules present in the tissue. However, its performance is severely hindered by low SNR, field inhomogeneities and overlapping spectra of metabolites, which affect the quantification of metabolites. Consequently, MRS is rarely included in routine clinical protocols and has not been proven in multi-institutional trials. This work proposes an alternative approach, where instead of quantifying metabolites' concentration, deep learning (DL) is used to model the complex nonlinear relationship between diseases and their spectroscopic metabolic fingerprint (pattern). DL requires large training datasets, acquired (ideally) with the same protocol/scanner, which are very rarely available. To overcome this limitation, a novel method is proposed that can quantum mechanically synthesise MRS data for any scanner/acquisition protocol. The proposed methodology is applied to the challenging clinical problem of differentiating metastasis from glioblastoma brain tumours on data acquired across multiple institutions. DL algorithms were trained on the augmented synthetic spectra and tested on two independent datasets acquired by different scanners, achieving a receiver operating characteristic area under the curve of up to 0.96 and 0.97, respectively.

KEYWORDS
deep learning, magnetic resonance spectroscopy, metabolic fingerprint, quantum mechanical spectroscopy, synthetic data

1 | INTRODUCTION

Human metabolism plays an important role in health conditions and can depict underlying disease mechanisms.[1] Identifying abnormal levels of certain metabolites can be useful for both diagnosis and treatment of diseases.[2] Magnetic resonance spectroscopy (MRS) can noninvasively evaluate metabolic activity and is used to diagnose metabolic disorders. MRS studies rely on the quantification of metabolic activity linked to a certain disease to detect, characterise and manage treatment. However, the quantification of the metabolites' concentration is significantly affected by limitations in the acquisition such as low signal-to-noise ratio (SNR), field inhomogeneities, as well as overlapping spectra of the metabolites. Consequently, it is not easy to depict the contribution in the MR signal of individual metabolites and accurately quantify their presence, particularly

Abbreviations used:: ANN, artificial neural network; AUC, area under the curve; CNN, convolutional neural network; LTE, long echo time; MRS, magnetic resonance spectroscopy; ROC, receiver operating characteristic curve; SNR, signal-to-noise ratio; STE, short echo time; SVM, support vector machine; TE, echo time; TR, repetition time.

[Corrections added on 25 January 2021, after first online publication: The graphical abstract and the short abstract were replaced.]

when their concentration is low relative to the sensitivity of MRS. High field MR scanners improve SNR,[3] but suffer more from field inhomogeneities, which need to be properly accounted for.

This work aims to quantum mechanically model synthetic MRS spectra that will be used to train machine learning algorithms to classify diseases based on the pattern of acquired in vivo spectra. The first underlying hypothesis is that diseases have a unique spectroscopy (metabolic) fingerprint[4] and their classification can be achieved based on the pattern of the spectra rather than the quantification of metabolic activity.[5] Being able to generate an abundance of synthetic data would enable the use of deep learning algorithms such as convolutional neural networks (CNNs).[6] Deep learning, according to the circuit complexity theory, can efficiently fit complex multivariate functions, and with the addition of a convolutional layer (as in a CNN) is suitable for pattern recognition within spectra. To date, the training of such complex models is not feasible as they require large training datasets that are very rarely available. Further, MR centres use different MR scanners and different acquisition protocols; hence, it is not feasible to combine data from multiple centres to train a predictive model. The second hypothesis behind this work is that deep learning models, trained on augmented synthetic MRS spectra, could be applied across different centres that use any scanner/acquisition protocol, significantly improving the diagnostic value of MRS. This is a major potential benefit of using synthetic MR spectra as they can be modelled for any scanner and for any acquisition protocol. Augmented techniques are also important to increase the amount of data and enable the use of machine learning algorithms.

The proposed methodology will be used to address the clinical problem of discriminating metastasis from glioblastoma brain cancers, based on their metabolic fingerprint as depicted by in vivo 1H MRS, which cannot be addressed with conventional imaging. Metastatic cancers are the most common brain tumours and are mostly caused by haematogenous spread originating from lung, breast, skin or colon cancers.[7] Primary glioblastoma multiforme (also called glioblastoma) is a fast-growing glioma that develops glial cells that often spread into nearby brain tissue. Glioblastoma is usually found in the cerebral hemispheres, and according to the American Association of Neurological Surgeons, it typically results in death 15 months after its diagnosis. Pretreatment discrimination of metastatic from glioblastoma brain tumours is a very important clinical problem as they have substantially different clinical management.[8] Metastatic brain cancers are commonly treated with radiation and stereotactic radiosurgery.[9] Glioblastoma's standard treatment includes intracranial surgery followed by radiotherapy with concomitant and adjuvant chemotherapy with temozolomide.[10] Metastasis and glioblastomas have a very similar MR imaging phenotype, appearing as solitary brain tumours surrounded by oedema, which makes them hard to differentiate based on conventional imaging. A range of advanced imaging methods such as

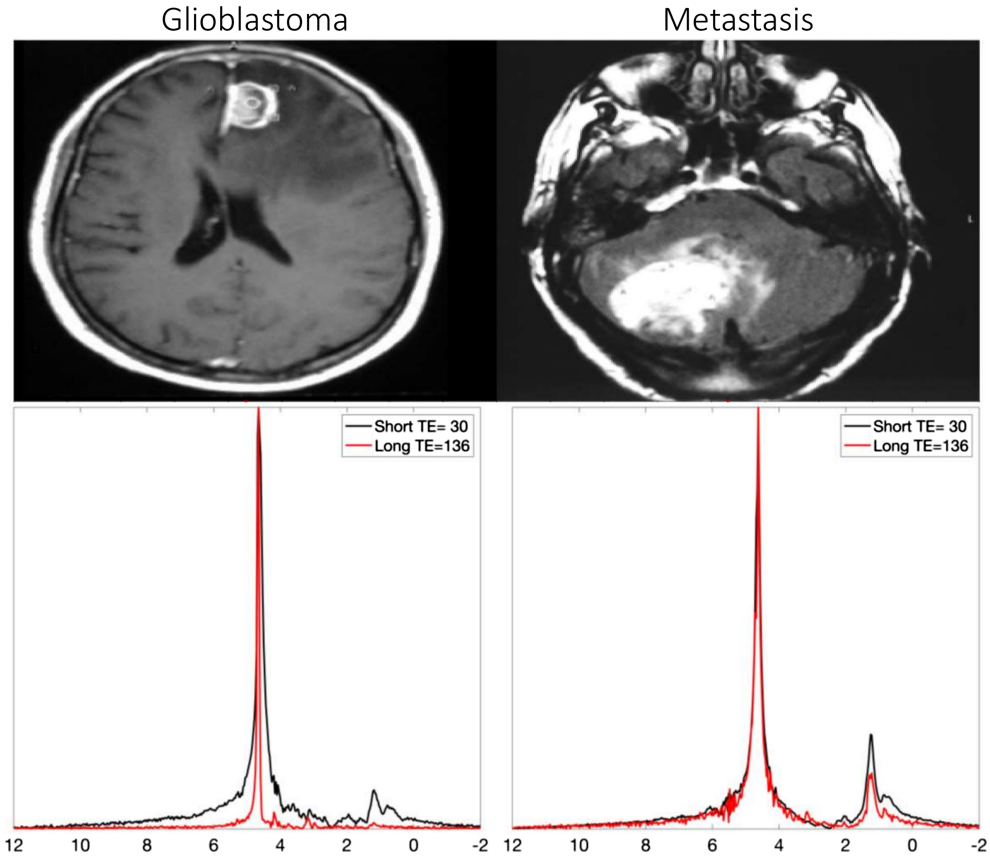

FIGURE 1 Metastasis and glioblastoma cases from the INTERPRET project. A T1-weighted postgadolinium image and MR spectra based on STE and LTE single-voxel sequences are shown per case

diffusion imaging,[11–13] dynamic contrast-enhanced imaging,[14] MRS[15,16] and radiomics[17–19] have been reported as showing promising results. However, with the exception of the study conducted by Velido et al.[16], these are single-centre studies, commonly performed on a single scanner, and have not been tested on independent patient populations. The proposed methodology used data acquired across multiple institutions.[20–24] CNN models were trained on the augmented synthetic spectra and tested on two independent datasets acquired by different scanners.

2 | MATERIALS AND METHODS

This work is based on the INTERPRET single-voxel multicentre dataset,[20–24] where each patient went through two different sequences using a short (STE) and a long echo time (LTE). STE spectra include more metabolites than LTE, but LTE key metabolites like creatine and choline have higher SNRs.[25] The two single-voxel sequences used in the INTERPRET dataset were Point RESolved Spectroscopy (PRESS)[26] and the stimulated echo acquisition mode (STEAM)[27] sequence. Figure 1 shows an example of a metastasis and a glioblastoma brain tumour acquired in the INTERPRET dataset with LTE and STE. Augmented synthetic spectra were modelled for every sequence based on a quantum mechanical model then were used to train 1D CNN models. The trained models were then tested on independently acquired spectra. Figure 2 outlines the methodology followed.

2.1 | Patient population and acquisition protocol

The INTERPRET study[20–24] consists of data from eight different clinical centres using three different 1.5 T MR scanners, namely, GE Signa Advantage and LX CV/i 1.5 T, Philips NT and ACS NT 1.5 T, and Siemens Vision 1.5 T. This work included 112 patients from the INTERPRET database with histopathologically confirmed glioblastoma (81 cases) and metastasis (31 cases). Each patient had single-voxel spectra acquired with both a long (using PRESS) and a short (using either PRESS or STEAM) echo time from a nodular region of the tumour where the biopsy was taken. Table 1 summarises the details of the acquisition and patient population.

2.2 | Quantum mechanically modelled MR spectra

In classical MR an ensemble of noninteracting spins (like protons in the hydrogen atoms in water molecules) are represented as a magnetisation vector **M**(t). The magnetisation vector is perturbed by a magnetic field **B1**(t) pulse, and once the **B1**(t) pulse stops, the magnetisation **M**(t) undergoes several processes such as T1 and T2 relaxation as it returns to equilibrium. The magnetic field **B**(t) may include the following: (i) offsets with respect to the main static magnetic field **B0**; (ii) radiofrequency (RF) pulses creating an additional magnetic field **B1**, which by the RF coil design can be complex, spatially dependent, and a superposition of multiple **B1** fields; and (iii) magnetic field gradients **G**(t) for spatial

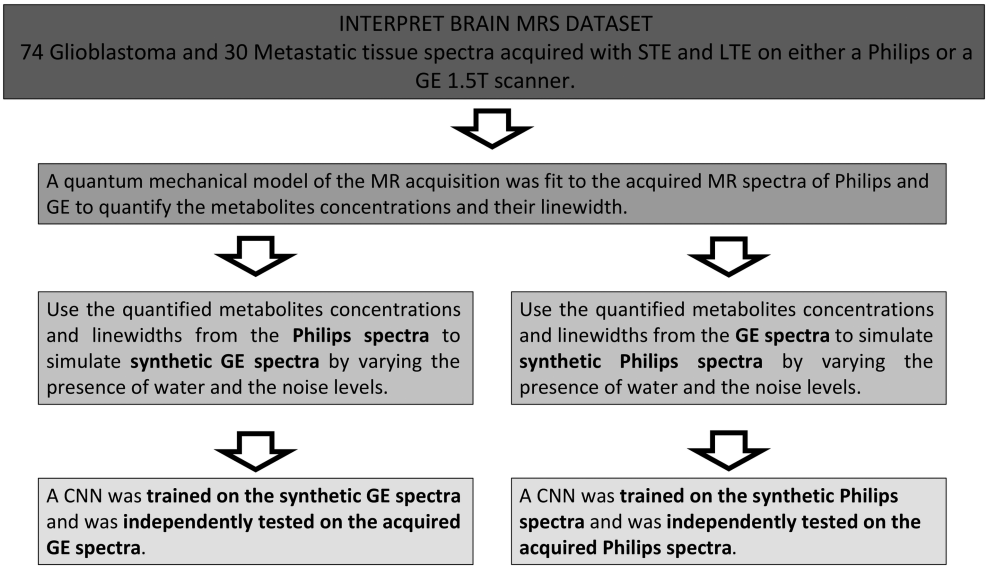

FIGURE 2 Flowchart summarising the methodology

TABLE 1 Description of the different acquisition protocols and patient population. The Siemens dataset was highly imbalanced and had a low number of samples, hence it was not included in this work

	Philips (48 patients)		GE				Siemens	
Volume (cm^3)	4–8		4–8				4–8	
N spectrum points	512		2048				1024	
Bandwidth (Hz)	1,000		2500				1000	
TR (ms)	2000		2000				2000	
Sequence	PRESS	PRESS	STEAM	STEAM	PRESS	PRESS	STEAM	PRESS
TE (ms)	30	136	20	30	30	135	20	136
Glioblastoma cases	32	32	30	8	5	43	6	6
Metastasis cases	16	16	4	3	7	14	1	1

encoding. The Bloch formulation can be summarised as $\frac{dM}{dt} = -\gamma B(t) \times M$, where $B(t) = B0 + B1(t) + r \cdot G(t)$. The classical model is sufficient to describe two state systems such as water molecules,[28] but metabolites with interacting spins require a quantum mechanical description.[29]

A quantum mechanical model based on the density matrix formalism was implemented in Python, similar to the implementation of Simpson et al.[30] Specifically, the time evolution of the density matrix, $\rho(t)$, is described by the Liouville-von Neumann equation $\frac{d\rho}{dt} = -i[H(t),\rho]$, which for a time-dependent Hamiltonian can be solved as $\rho(t) = e^{-itH}\rho(0)e^{iH}$. The total spin Hamiltonian H reflects internal nuclear spin energy interactions (chemical shifts H_{CS}, dipole–dipole couplings H_{DD}, J-couplings H_J, and quadrupolar coupling H_Q interactions) and RF pulses H_{RF}, $H_{TOT}(t) = H_{CS}(t) + H_J(t) + H_{DD}(t) + H_Z(t) + H_Q(t) + H_{RF}(t)$. This work assumes that the motion of molecules is fast, isotropic and hence that $\omega_0\tau_C << 1$, where ω_0 is the resonant frequency and τ_C is the correlation time, and as a result the dipole–dipole couplings H_{DD} and the quadrupolar couplings H_Q average to zero. 1H MRS is a two spin-½ nuclei AB system (with angular momentum operators I and S, respectively) and without RF pulses is represented by $H = \omega_A I_z + \omega_B S_z + J_{AB}(I \cdot S)$, where J_{AB} is the coupling constant and ω_A,ω_B are the chemical shifts of spin A and B, respectively. The density matrix formalism is described in more detail by Mulkern and Bowers.[29]

Both STEAM and PRESS have been quantum mechanically modelled using the density matrix formalism. Excitation, refocusing pulses and signal readout were implemented as described by Simpson et al.[30] and Mulkern et al.[31,32] Aside from H_2O, 24 different metabolites were modelled, namely, alanine, aspartate, phospho-choline, creatine, phospho-creatine, citrate, gamma-aminobutyric acid, glutamine, glutamate, glutathione, glycine, myo-inositol, lactate, N-acetyl aspartate, scyllo-inositol, taurine, ascorbate, beta-hydroxybutyrate, 2-hydroxyglutarate, glucose, N-acetyl aspartyl glutamate, glycero-phosphocholine, phosphoryl ethanolamine and serine. The resulting free induction decay (FID) signals per metabolite were added and the sum was Fourier-transformed to frequency space to give the modelled spectrum. The model depended on a series of acquisition parameters, namely, field strength, the number of spectral points, spectral width, echo times, repetition time, type of sequence and parameters related to the presence of metabolites (ie, the number of atoms [*Nat*], *T2** per metabolite).

2.3 | Mathematical optimisation

The modelled MR spectra were fitted to the acquired spectra as a function of the number of atoms and the spin–spin relaxation time *T2** of each metabolite. The *Nat* dictate the presence of the metabolite in the region of interest. The *T2** affects the full width half maximum of the resonance lineshape $\Delta\nu_{FWHM} = 1/(\pi T2^*)$ and consequently both the frequency resolution and the SNR of the spectra. *T2** is an intrinsic molecular property but is also affected by magnetic field irregularities. Consequently, its value cannot be accurately known in advance and was included as a parameter in the fitting process.

Given the large number of unknowns (*Nat*, *T2** × 25 molecules = 50 unknowns), a novel optimisation algorithm was implemented that hierarchically fitted the signal from each molecule to the acquired spectra (always starting from H_2O, which dominates the signal) until all molecules were included. Fitting was performed by minimising the L1-norm, using the simplex algorithm, which is an iterative procedure that does not need information regarding the derivative of the function under consideration. The algorithm creates a 'random' simplex of n + 1 points, where n is the number of the model parameters that need to be estimated. The simplex changes iteratively by reflection, expansion and contraction steps, until it finds the model parameters that minimise the given likelihood function. The constrained variation of the simplex algorithm[33] was used for all tested models; an L1-norm was employed in the likelihood function to improve robustness. The simplex algorithm is particularly effective for cases like this one where the gradient of the likelihood functions is not easy to calculate. The simplex algorithm was chosen because it performed better than other curve fitting algorithms, namely, the Levenberg-Marquardt,[34] the trust-region-reflective[35] and the adaptive learning rate stochastic optimisation (ADAM).[36]

An update on metabolites *Nat* and *T2** was accepted only if it improved the goodness of fit (*gof*) measured by two-sample Kolmogorov-Smirnov test (ks-test). This process was repeated until the signal from all metabolites was updated 100 times. The proposed hierarchical simplex algorithm is summarised in Table 2.

2.4 | Augmented (synthetic) MR spectra

The *Nat* and *T2** of the different metabolites were quantified for the different spectra of glioblastoma and metastasis brain tumours acquired using the Philips and the GE scanner. The noise levels of the acquired spectra were also estimated by the median absolute deviation technique proposed by Dikaios et al.[37] Based on the quantified {*Nat*, *T2**} parameters of each metabolite from the Philips spectra, augmented synthetic GE spectra were generated by varying the *Nat* for the H_2O molecules and by adding different levels of noise. Specifically, for each augmented GE spectra:

(1) 10 different values of N_{at}^{H20} were used between the range of minimum and maximum quantified N_{at}^{H20} from the Philips spectra.

(2) 10 different noise levels were added (as Gaussian noise) to the modelled FID between the range of minimum and maximum quantified noise levels from the Philips spectra.Knowing the {*Nat*, *T2**} parameters for all metabolites from the Philips spectra and the GE acquisition parameters for the different sequences, synthetic GE spectra were modelled. This process resulted in 48 × 10 × 10 augmented LTE (TE = 135 ms) PRESS GE spectra, and 48 × 10 × 10 = 4800 augmented STE (TE = 30 ms) PRESS GE spectra.

A similar process was followed to generate augmented Philips data based on the quantified {*Nat*, *T2**} parameters of all metabolites from the GE spectra. Similarly, 57 × 10 × 10 = 5700 LTE (TE = 136 ms) PRESS Philips spectra and 57 × 10 × 10 = 5700 STE (TE = 30 ms) PRESS Philips spectra were modelled.

Synthetic data were not generated based on the Siemens scanner as the number of samples was too low and imbalanced to draw any conclusions (six glioblastomas and one metastasis).

2.5 | Predictive modelling based on training with augmented acquired spectra

Support vector machine (SVM), artificial neural network (ANN) and 1D CNN machine learning models were trained on the GE or the Philips spectra and independently tested on the Philips or the GE spectra, respectively. To maximise the performance of the models, for every spectrum used in the training, in addition to the measured data, two different noise levels were added (plus 10% and 20% of the estimated noise), which significantly increased the number of spectra used in the training. Adding noise is a fundamental tool for data augmentation.[38] This resulted in 3 × 57 GE spectra and 3 × 48 Philips spectra and was found to be beneficial in the performance of the trained models. Hyperparameters and/or model parameters were optimised using a grid search approach for each classifier as part of the training. The optimal parameters were selected based on the best mean cross-validation accuracy. Ten-fold cross validation was performed for all algorithms.

SVMs were run with regularisation parameter C = 0.1 and kernel coefficient gamma = 100. ANNs were run with two hidden layers of 200 neurons each, a total training epoch of 200, learning rate (0.001, 0.01, 0.1, 0.2), weight decay 0.0005 and a training batch of 32. 1D CNNs were configured with three convolutional layers with the number of filters equal to 24, 12 and 8, respectively, followed by a ReLU layer and 1 × 3 Max pooling layer. There were two fully connected layers with 200 fully connected neurons each with a ReLU layer in between, followed by a softmax and classification layer. The ADAM method was used for optimisation.

TABLE 2 Hierarchical simplex algorithm to fit the quantum mechanical MRS model Φ as a function of *Nat*, *T2** to the acquired spectra (**y**)

Initialise *Nat*, *T2 parameters for each metabolite, m**
Normalise the acquired spectra, *y*
gof0 = *ks-test*(**Φ**(*Nat*, *T2**), **y**)
While *i* < 100:
Randomly permute the order metabolites *Nat*(*m*), *T2**(*m*) are updated, but always starting with H_2O
For each metabolite m:
Minimise $\|\mathbf{\Phi}(Nat, T2^*) - \mathbf{y}\|_1$ with respect to *Nat*(***m***), *T2**(***m***) using the simplex
gof = *ks - test*(Φ(*Nat*, *T2**), *y*)
If *gof* < *gof0*
Update *Nat*(***m***), *T2**(***m***), *gof0*
end
end

2.6 | Predictive modelling based on training with augmented synthetic spectra

SVM, ANN and 1D CNN models were trained on the augmented synthetic GE/Philips spectra and were tested on the independently acquired GE/Philips spectra. All algorithms were optimised using grid search on the training augmented synthetic data, as already described.

SVMs were run with regularisation parameter C = 0.1 and kernel coefficient gamma = 100. ANNs were run with four hidden layers of 400 neurons each, total training epoch 200, learning rate (0.001, 0.01, 0.1, 0.2), weight decay 0.0005 and a training batch of 256. 1D CNNs were configured with four convolutional layers with numbers of filter 56, 24, 12 and 8, respectively, followed by a ReLU layer and 1×3 Max pooling layer. There were two fully connected layers with 300 fully connected neurons each with a ReLU layer in between, followed by a softmax and classification layer. The ADAM method was used for optimisation. Specifically, the machine learning algorithms were trained on six different datasets:

1. 4800 synthetic GE STE spectra.
2. 4800 synthetic GE LTE spectra.
3. 4800 synthetic GE concatenated STE and LTE spectra in a 1D sequence.
4. 5700 synthetic PHILIPS STE spectra.
5. 5700 synthetic PHILIPS LTE spectra.
6. 5700 synthetic PHILIPS concatenated STE and LTE spectra in a 1D sequence.

3 | RESULTS

3.1 | Quantum mechanically modelling the MRS signal of different acquisition protocols

Typically, the MR signal is modelled by the Bloch equations that represent an ensemble of noninteracting spins (like protons in the hydrogen atoms in water molecules). Water molecules in nuclear magnetic resonance can be perceived as two state systems (qubit) and, according to Feynman et al.,[28] can be fully described by classical MR. In MRS, however, certain metabolites have interacting spins, hence the signal needs to be quantum mechanically modelled using density matrix formalism (eg, signal from metabolites with a strongly coupled AB system such as citrate).

To showcase why the Bloch equations cannot accurately model the MRS signal and why a quantum mechanical model needs to be used, a phantom was built where aqueous solutions of citrate, creatine and choline were mixed in a plastic container and single-voxel PRESS spectra were obtained with flip angle 90°, echo times TE1 = 14.4 ms and TE2 = 15.6 ms, and a repetition time (TR) of 1410 ms on a Siemens 3 T scanner. The sequence was modelled with the Bloch and the Liouville-von Neumann equations, and the models were fitted to the pulse-acquired spectra. The two models displayed a similar performance on the creatine and the choline peaks, but the Bloch equation clearly failed to model the citrate peak (Figure 3).

3.2 | Mathematical optimisation

The quantum mechanical model was fitted to the acquired PRESS and STEAM spectra in the Philips and GE scanners to estimate the *Nat* and the linewidth (*T2**) of each metabolite, as described in the Materials and Methods section. A new hierarchical simplex algorithm was implemented to

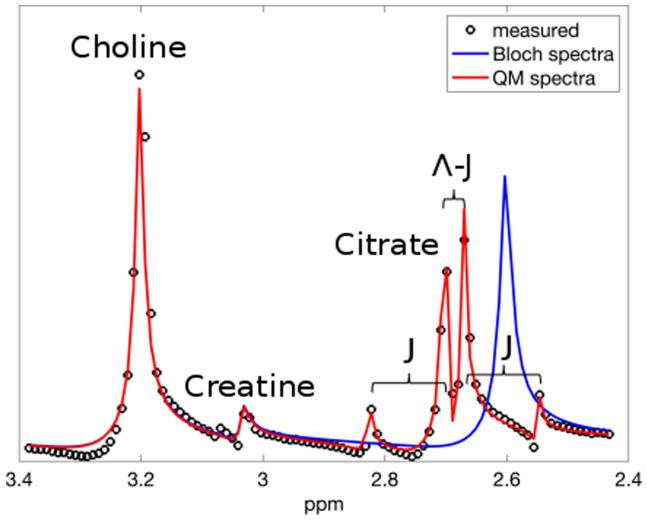

FIGURE 3 MR spectra in a pulse-acquired experiment of a phantom containing 30 mM of citrate, 4 mM of creatine and 15 mM of choline. Citrate ($^{-}$OOC-CH2-C (OH)(COO^{-})-CH2-COO^{-}) consists of two pairs of strongly coupled methylene protons. Consequently, J-coupling gives rise to line splitting, where a tall central doublet is split into two smaller peaks on either side. *J* denotes the strength of the J-coupling in Hz, $\Lambda = \sqrt{\Delta\nu^2 + J^2}$, where $\Delta\nu$ is the Larmor frequency difference between the two inequivalent protons

fit the spectra and was evaluated against publicly available software (ie, LCModel[39] and the AMARES package[40] in jMRUI).[41] LCModel also uses the density matrix formalism by Mulkern and Bowers.[29] The optimisations were compared in terms of *gof* using the ks-test. The proposed fitting method yielded a median ks-test across all samples of 0.278 (interquartile range 0.041) compared with 0.287 (interquartile range 0.041) for the LCModel and 0.317 (interquartile range 0.043) for Amares. Figure 4 demonstrates examples of the model fitted using the hierarchical simplex algorithm to LTE and STE spectra for both the Philips and GE scanners.

3.3 | Predictive modelling based on training with augmented acquired spectra

Three different machine learning algorithms were trained on the GE or Philips spectra then were independently tested on the Philips or GE spectra, respectively, as described in the Materials and Methods section. Namely, the investigated algorithms were SVMs, ANNs and 1D CNNs. The

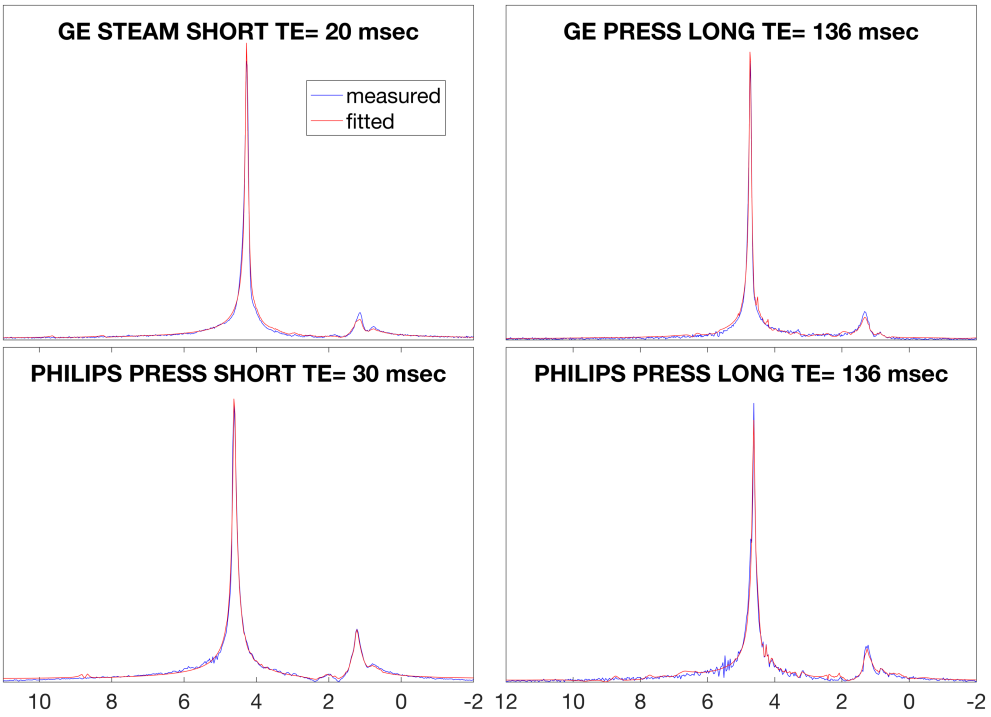

FIGURE 4 Examples of STE and LTE spectra fitted with the quantum mechanical model for metastasis brain tumours

TABLE 3 Performance of different machine learning algorithms trained on the noise-augmented spectra to discriminate metastasis from glioblastoma and independently tested on nonaugmented spectra. The assessment is based on accuracy and ROC-AUC (95% confidence interval)

Trained on 3 × 57 GE spectra and tested on 48 Philips spectra				
		SVM	ANN	CNN
ACCURACY	STE	0.67	0.73	0.77
	LTE	0.69	0.75	0.77
	STE + LTE	0.75	0.75	0.83
AUC	STE	NA	0.79 (0.70–0.86)	0.84 (0.80–0.91)
	LTE	NA	0.82 (0.75–0.87)	0.85 (0.77–0.91)
	STE + LTE	NA	0.86 (0.78–0.91)	0.88 (0.77–0.93)
Trained on 3 × 48 Philips spectra and tested on 57 GE spectra				
		SVM	ANN	CNN
ACCURACY	STE	0.60	0.67	0.72
	LTE	0.60	0.70	0.72
	STE + LTE	0.61	0.70	0.74
AUC	STE	NA	0.76 (0.63–0.88)	0.80 (0.74–0.89)
	LTE	NA	0.76 (0.65–0.88)	0.80 (0.75–0.90)
	STE + LTE	NA	0.80 (0.73–0.90)	0.83 (0.77–0.91)

spectra used for training were augmented by adding two different (Gaussian) noise levels to the FID MR signal before Fourier-transforming it to spectra. Adding noise is a fundamental tool for data augmentation as it provides more training data and improves the generalisation power of the trained model.[38] Table 3 summarises the performance of the machine learning algorithms trained on the noise-augmented data in terms of the receiver operating characteristic (ROC) area under the curve (AUC) and accuracy using a 50% threshold. Note that each machine learning algorithm was trained on STE, LTE and STE + LTE spectra. Noise augmentation improved the performance (in terms of ROC-AUC) of the machine learning algorithms on the spectra used for testing (nonaugmented) by up to 3% for SVMs, 8% for ANNs, and 7% for 1D CNNs. Further augmenting the training dataset by adding more noise levels was attempted but this was not beneficial.

3.4 | Predictive modelling based on training with augmented synthetic spectra

The *Nat* and the respective linewidths (*T2**) of each metabolite were estimated for all acquired spectra. Synthetic GE spectra were quantum mechanically modelled based on the estimated *Nat*, *T2** from the Philips spectra and using the GE acquisition parameters. The synthetic GE spectra were augmented by varying *Nat* for H_2O molecules and the level of noise, as described in the Materials and Methods section. The MRS spectrum is heavily dominated by H_2O molecules that have approximately 10,000 times higher concentrations than those of the metabolites of interest. The H_2O peak in MRS spectra can in general vary as a function of the water presence in the tumour as well as partial volume with adipose within the voxel. This may not be the case for GBM spectra when signals come from necrotic tissues. The suggested augmentation by varying the H_2O *Nat* can lead to synthesising more spectra without affecting the metabolic fingerprint of the tissue.

An equivalent process was followed to generate augmented synthetic Philips spectra, again by varying *Nat* for H_2O molecules and the level of noise. The augmented synthetic GE/Philips spectra were used to train SVM, ANN and 1D CNN models and were tested on the acquired Philips/GE spectra, respectively. Table 4 shows the diagnostic performance of each machine learning algorithm trained on augmented synthetic spectra and tested on independently acquired (nonaugmented) spectra in terms of ROC-AUC and accuracy. Figure 5 shows the performance of 1D CNN models trained on STE, LTE and STE + LTE spectra.

4 | DISCUSSION

MRS is a valuable noninvasive diagnostic imaging tool that can depict metabolic disorders. Despite its potential, its performance is severely hindered by low SNR, field inhomogeneities and overlapping spectra of the different metabolites that limit its ability to quantify metabolic activity. Consequently, MRS is very rarely included in clinical trials[42,43] and has not yet been proven in multi-institutional trials. The main objective of this study was to improve the diagnostic value of MRS by understanding the complex nonlinear relationship between the metabolic spectroscopy fingerprint and diseases. There were four main contributions.

Firstly, a quantum mechanical model implemented in Python that can model the MRS free induction decay signal based on the acquisition parameters and the *Nat*/*T2** of each metabolite.

TABLE 4 Performance of different machine learning algorithms trained on augmented synthetic spectra to discriminate metastasis from glioblastoma and independently tested on nonaugmented spectra. The assessment is based on accuracy and ROC-AUC (95% confidence interval)

Trained on augmented synthetic 4800 Philips spectra and tested on 48 Philips-acquired spectra				
		SVM	ANN	CNN
ACCURACY	STE	0.71	0.85	0.88
	LTE	0.71	0.85	0.88
	STE + LTE	0.75	0.88	0.92
AUC	STE	NA	0.89 (0.81–0.93)	0.94 (0.87–0.98)
	LTE	NA	0.92 (0.82–0.95)	0.95 (0.89–1.00)
	STE + LTE	NA	0.93 (0.82–0.95)	0.96 (0.90–1.00)
Trained on augmented synthetic 5700 GE spectra and tested on 57 GE-acquired spectra				
		SVM	ANN	CNN
ACCURACY	STE	0.74	0.88	0.88
	LTE	0.77	0.88	0.93
	STE + LTE	0.77	0.90	0.93
AUC	STE	NA	0.93 (0.89–0.98)	0.91 (0.88–0.99)
	LTE	NA	0.94 (0.87–0.99)	0.95 (0.90–1.00)
	STE + LTE	NA	0.94 (0.90–0.99)	0.97 (0.89–1.00)

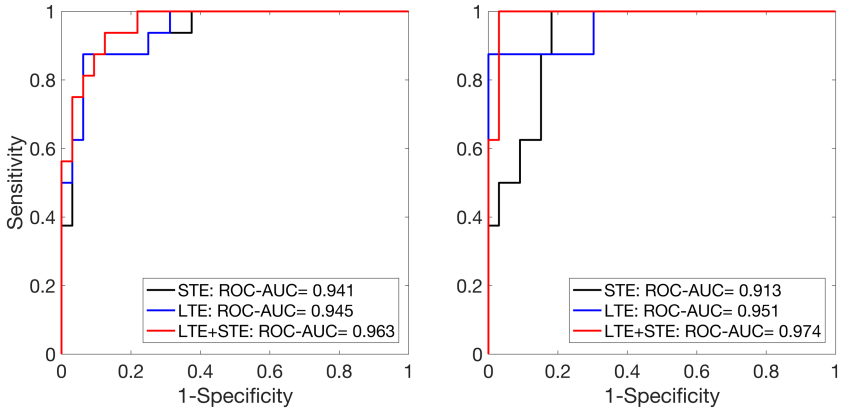

FIGURE 5 Receiver operating characteristic (ROC) curve shown for the 1D CNN models (STE, LTE, and STE + LTE) tested on the single-voxel spectra from the INTERPRET project. Left: 1D CNN model trained on augmented synthetic Philips spectra and tested on acquired Philips spectra. Right: 1D CNN model trained on augmented synthetic GE spectra and tested on acquired GE spectra

Secondly, a novel optimisation method to fit the quantum mechanical model to the measured spectra and estimate *Nat*/*T2** for each metabolite. The proposed hierarchical simplex method, as well as accurately matching the acquired spectra, is automatic and requires no user input.

Thirdly, the generation of synthetic spectra for any MR scanner by varying the acquisition parameters and using the estimated *Nat*/*T2** of each metabolite. The synthetic data were then augmented by using a novel method suitable to MRS. Specifically, the variation in the *Nat* of water molecules is suggested, which accounts for variations in the water content of tumours while not affecting the metabolic fingerprint. This, in combination with varying the noise levels, increased the number of spectra 100 times, enabling the use of deep learning algorithms.

Fourthly, the training of machine learning algorithms based on augmented synthetic data, including the 1D CNN, which is particularly suitable for pattern recognition problems. Even more importantly, the synthetic spectra used for training were modelled based on parameters estimated from a different scanner than the one that was used to test the performance of the 1D CNN models. This is particularly important because different centres have different scanners and often use different acquisition protocols. A methodology that can generate augmented synthetic data for any scanner/acquisition protocol then use them to train an accurate diagnostic model would be extremely beneficial, especially when large datasets are not available.

The proposed methodology could in theory be followed for any classification problem involving MRS, but in this work, it was used to differentiate metastasis from glioblastoma brain tumours based on the INTERPRET single-voxel MRS multicentre dataset. This is a critical and very challenging clinical problem, as both tumour types share the same clinical symptoms and cannot be discriminated with MR imaging. MRS analysis has shown promise in differentiating brain tumours; however, the INTERPRET multicentre study recognised the difficulty in differentiating glioblastomas from metastasis.[21] A more recent study[16] of the same INTERPRET dataset reported an accuracy of up to 85% and an ROC-AUC of up to 0.86 on a testing dataset of 30 glioblastomas and 10 metastases acquired across three different centres (and three different MR scanners: Siemens, GE and Philips). The proposed 1D-CNN models trained on augmented synthetic data had an accuracy (Table 2) of up to 93% and an ROC-AUC of up to 0.97 when evaluated on two larger (48 and 57 patients, respectively), independent (tested on spectra from a different scanner than the one used to synthesise the training data) datasets from the same INTERPRET study. To investigate the benefits of using synthetic data for training, machine learning algorithms were also trained on the acquired Philips/GE augmented with noise addition and were independently tested on the nonaugmented GE/Philips spectra, respectively (Table 1). The reported accuracies/ROC-AUC were consistently lower than for the machine learning algorithms trained on the synthetic data.

In agreement with other studies, LTE spectra could better discriminate glioblastomas from metastasis compared with STE. Although LTE single-voxel spectroscopy depicts fewer metabolites, it shows more distinct peaks, which seems to benefit the pattern recognition. Further, the combination of STE and LTE spectra was found to be consistently better compared with using only LTE or STE.

The current study has some potential limitations. Firstly, the MRS spectra available from the INTERPRET study used to generate the synthetic data and test the performance of the 1D CNN models were acquired solely on 1.5 T scanners. Secondly, the proposed methodology needs to be applied to other clinical problems, such as early detection of brain tumours and differentiating more brain tumour types. Future work will involve testing the 1D CNN models on MRS data from 3 T scanners[3] and applying the proposed methodology to other clinical problems.

To conclude, the findings support the two main hypotheses that diseases can have a distinct metabolic (and spectroscopy) fingerprint, and that augmented synthetic spectra can be modelled for any MR acquisition protocol/scanner to train diagnostic models. Generating large augmented synthetic data is particularly important, especially when deep learning is used for training the diagnostic models. In turn, enabling the use of deep learning methods by using augmented synthetic data can significantly improve our ability to classify diseases and resolve challenging clinical problems such as the differentiation of metastasis from glioblastoma. Finally, the proposed methodology could be applied across different

institutions synthesising MRS data, even when acquisition protocols change, and can also train deep learning algorithms, even when large datasets are not available.

ACKNOWLEDGEMENTS

The author is grateful to Dr Margarida Julià-Sapé for her invaluable help with the INTERPRET dataset and to Prof. John Griffith for suggesting the dataset. Nikolaos Dikaios was supported by the Royal Society (INF\R1\191030).

FUNDING INFORMATION

Nikolaos Dikaios was supported by the Royal Society (INF\R1\191030).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

ORCID

Nikolaos Dikaios https://orcid.org/0000-0001-9865-0260

REFERENCES

1. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17(4):448-453.
2. Lee D-S, Park J, Kay K, Christakis NA, Oltvai Z, Barabási A-L. The implications of human metabolic network topology for disease comorbidity. *Proc Natl Acad Sci*. 2008;105(29):9880-9885.
3. Fuster-Garcia E, Navarro C, Vicente J, et al. Compatibility between 3T 1H SV-MRS data and automatic brain tumour diagnosis support systems based on databases of 1.5T 1H SV-MRS spectra. *Magma*. 2011;24(1):35-42.
4. Kosmides AK, Kamisoglu K, Calvano SE, Corbett SA, Androulakis IP. Metabolomic fingerprinting: challenges and opportunities. *Crit Rev Biomed Eng*. 2013;41(3):205-221.
5. Ortega-Martorell S, Julià-Sapé M, Lisboa P, Arús C. Pattern Recognition Analysis of MR Spectra. In: Griffiths J, Bottomley P, eds. *Handbook of in vivo Magnetic Resonance Spectroscopy*. Hoboken, New Jersey: John Wiley & Sons; 2016.
6. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86(11):2278-2324.
7. Sawaya R, Bindal RK, Lang FF, Abi-Said D. Metastatic brain tumors. In: Kaye AH, Laws ER Jr, eds. *Brain Tumours: an encyclopedia approach*. London, UK: Harcourt Publishers Limited; 2001:999-1026.
8. Neska-Matuszewska M, Bladowska J, Sąsiadek M, Zimny A. Differentiation of glioblastoma multiforme, metastases and primary central nervous system lymphomas using multiparametric perfusion and diffusion MR imaging of a tumor core and a peritumoral zone-Searching for a practical approach. *PLoS ONE*. 2018;13(1):1913-1941, e0191341.
9. Platta CS, Khuntia D, Mehta MP, Suh JH. Current treatment strategies for brain metastasis and complications from therapeutic techniques: a review of current literature. *Am J Clin Oncol*. 2010;33(4):398-407.
10. Minniti G, Muni R, Lanzetta G, Marchetti P, Enrici RM. Chemotherapy for glioblastoma: current treatment and future perspectives for cytotoxic and targeted agents. *Anticancer Res*. 2009;29(12):5171-5184.
11. El-Serougy LG, Razek AAKA, Mousa AE, Eldawoody HAF, El-Morsy AE-ME. Differentiation between high-grade gliomas and metastatic brain tumors using diffusion tensor imaging metrics. *Egypt J Radiol Nucl Med*. 2015;46(4):1099-1104.
12. Byrnes TJ, Barrick TR, Bell BA, Clark CA. Diffusion tensor imaging discriminates between glioblastoma and cerebral metastases in vivo. *NMR Biomed*. 2011;24(1):54-60.
13. Abdel Razek AAK, Talaat M, El-Serougy L, Abdelsalam M, Gaballa G. Differentiating glioblastomas from solitary brain metastases using arterial spin labeling perfusion- and diffusion tensor imaging-derived metrics. *World Neurosurg*. 2019;127:593-598.
14. Law M, Cha S, Knopp EA, Johnson G, Arnett J, Litt AW. High-grade gliomas and solitary metastases: Differentiation by using perfusion and proton spectroscopic MR imaging. *Radiology*. 2002;222(3):715-721.
15. Fan G, Sun B, Wu Z, Guo Q, Guo Y. In vivo single-voxel proton MR spectroscopy in the differentiation of high-grade gliomas and solitary metastases. *Clin Radiol*. 2004;5(9):77-85.17.
16. Vellido A, Romero E, Julià-Sapé M, et al. Robust discrimination of glioblastomas from metastatic brain tumors on the basis of single-voxel (1)H MRS. *NMR Biomed*. 2012;25(6):819-282.
17. Chen C, Ou X, Wang J, Guo W, Ma X. Radiomics-based machine learning in differentiation between glioblastoma and metastatic brain tumors. *Front Oncol*. 2019;9:806.
18. Artzi M, Bressler I, Ben BD. Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis. *J Magn Reson Imaging*. 2019 Aug;50(2):519-528.
19. Artzi M, Bressler I, Ben BD. Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis. *J Magn Reson Imaging*. 2019;50(2):519-528.
20. Tate AR, Underwood J, Acosta DM, et al. Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra. *NMR Biomed*. 2006;19(4):411-434.
21. García-Gómez JM, Luts J, Julià-Sapé M, et al. Multiproject-multicenter evaluation of automatic brain tumor classification by magnetic resonance spectroscopy. *Magma*. 2009;22(1):5-18.
22. Pérez-Ruiz A, Julià-Sapé M, Mercadal G, Olier I, Majós C, Arús C. The INTERPRET Decision-Support System version 3.0 for evaluation of magnetic resonance spectroscopy data from human brain tumours and other abnormal brain masses. *BMC Bioinformatics*. 2010;11(1):581-597.

23. Julià-Sapé M, Majós C, Camins À, et al. Multicentre evaluation of the INTERPRET decision support system 2.0 for brain tumour classification. *NMR Biomed*. 2014;27(9):1009-1036.
24. Julià-Sapé M, Griffiths JR, Tate AR, et al. Classification of brain tumours from MR spectra: the INTERPRET collaboration and its outcomes. *NMR Biomed*. 2016;28(12):1772-1787.
25. Bogdan AA, Kataeva GV, Khomenko JG, Ilves AG, Prakhova LN. Diagnostic value of short and long echo time in 1H-MRS for patients with multiple sclerosis. *Appl Magn Reson*. 2017;48(7):707-714.
26. Bottomley PA. Spatial localization in NMR spectroscopy in vivo. *Ann N Y Acad Sci*. 1987;508(1 Physiological):333-348.
27. Frahm J, Merboldt K-D, Hanicke W. Localized proton spectroscopy using stimulated echoes. *J Magn Reson*. 1987;72:502-508.
28. Feynman RP, Vernon FL Jr, Hellwarth RW. Geometrical representation of the Schrodinger equation for solving MASER problems. *J Appl Phys*. 1957;28(1):49-52.
29. Mulkern R, Bowers J. Density matrix calculations of AB spectra from multipulse sequences: quantum mechanics meets in vivo spectroscopy concepts. *Magn Reson*. 1994;6:1-23.
30. Simpson R, Devenyi GA, Jezzard P, Hennessy TJ, Near J. Advanced processing and simulation of MRS data using the FID appliance (FID-A)—an open source. *MATLAB-Based Toolkit Magnetic Resonance in Medicine*. 2017;77(1):23-33.
31. Mulkern RV, Bowers JL. Calculating spectral modulations of AB systems during PRESS acquisitions. *Magn Reson Med*. 1993;30(4):518-519.
32. Mulkern RV, Bowers JL, Peled S, Williamson DS. Density-matrix calculations of the 1.5 T citrate signal acquired with volume-localized STEAM sequences. *J Magn Reson B*. 1996;110(3):255-266.
33. Lagarias JC, Reeds JA, Wright MH, Wright PE. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J Optimization*. 1998;9(1):112-147.
34. Moré JJ. The Levenberg-Marquardt Algorithm: Implementation and Theory. Numerical Analysis. In: Watson GA, ed. *Lecture Notes in Mathematics 630*. Salmon Tower Building: Springer Verlag:1977, 105-1116.
35. Coleman TF, Li Y. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM J Optimization*. 1996;6(2):418-445.
36. Diederik PK, Jimmy B. ADAM: A method for stochastic optimization. ICLR; 2014 (arXiv:1412.6980)
37. Dikaios N, Punwani S, Hamy V, et al. Noise estimation from averaged diffusion weighted images: Can unbiased quantitative decay parameters assist cancer evaluation? *Magn Reson Med*. 2014;71(6):2105-2117.
38. Sáiz-Abajoa MJ, Mevik BH, Segtnan VH, Næs T. Ensemble methods and data augmentation by noise addition applied to the analysis of spectroscopic data. *Anal Chim Acta*. 2005;533(2):147-159.
39. Provencher SW. Automatic quantitation of localized in vivo 1H spectra with LCModel. *NMR Biomed*. 2001;14(4):260-264.
40. Vanhamme L, van den Boogaart A, Van Huffel S. Improved method for accurate and efficient quantification of MRS data with use of prior knowledge. *J Magn Reson*. 1997;129(1):35-43.
41. Naressi A, Couturier C, Devos J, et al. Java-based graphical user interface for the MRUI quantitation package. *Magn Reson Mater Phy*. 2001;12(2-3):141-152.
42. Barentsz JO, Weinreb JC, Verma S, et al. Synopsis of the PI-RADS v2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. *Eur Urol*. 2016;69(1):41-49.
43. Thust SC, Heiland S, Falini A, et al. Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice. *Eur Radiol*. 2018;28(8):3306-3317.

How to cite this article: Dikaios N. Deep learning magnetic resonance spectroscopy fingerprints of brain tumours using quantum mechanically synthesised data. *NMR in Biomedicine*. 2021;34:e4479. https://doi.org/10.1002/nbm.4479