Aus der Klinik und Poliklinik für Radiologie
Klinikum der Ludwig-Maximilians-Universität München
Direktor: Prof. Dr. med. Jens Ricke

Robust Evaluation of Contrast-Enhanced Imaging for Perfusion Quantification

Kumulative Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
an der Medizinischen Fakultät der
Ludwig–Maximilians–Universität zu München

vorgelegt von
Andreas Mittermeier
aus
Burghausen

2022

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Betreuer:	Prof. Dr. rer. nat. Michael Ingrisch
Zweitgutachter:	Prof. Dr. rer. nat. Guillaume Landry
Dekan:	Prof. Dr. Thomas Gudermann
Tag der mündlichen Prüfung:	10. Mai 2023

Eidesstattliche Versicherung

Mittermeier, Andreas

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Titel

Robust Evaluation of Contrast-Enhanced Imaging for Perfusion Quantification

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht habe und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 17. Mai 2023

Ort, Datum

Andreas Mittermeier

Unterschrift

Contents

List of Abbreviations

1D	One Dimensional
2CX	Two Compartment Exchange
2D	Two Dimensional
3D	Three Dimensional
4D	Four Dimensional
AIF	Arterial Input Function
ASPECTS	Alberta Stroke Program Early CT Score
AUC	Area Under the Curve
BTM	Bayesian Tofts Model
CA	Contrast Agent
CBF	Cerebral Blood Flow
CBV	Cerebral Blood Volume
CNN	Convolutional Neural Network
CT	Computed Tomography
CTP	Computed Tomography Perfusion
CV	Cross Validation
DCE	Dynamic Contrast-Enhanced
DL	Deep Learning
DRO	Digital Reference Object
EES	Extravascular Extracellular Space
FC	Fully Connected
FID	Free Induction Decay
FOV	Field Of View
HDI	Highest Density Intervals
HU	Houndsfield Units
IRF	Impulse Response Function
L-BFGS	Limited-memory Broyden–Fletcher–Goldfarb–Shanno
MCMC	Markov Chain Monte Carlo
MRI	Magnetic Resonance Imaging
MTT	Mean Transit Time
NACT	Neoadjuvant Chemotherapy
NLLS	Non-Linear Least Squares

NMR	Nuclear Magnetic Resonance
NUTS	No-U-Turn Algorithm
pCR	pathologic Complete Response
PK	Pharmacokinetic
PPC	Posterior Predictive Check
PS	Permeability-Surface Area Product
QIBA	Quantitative Imaging Biomarkers Alliance
QIN	Quantitative Imaging Network
RF	Radiofrequency
RMSE	Root-Mean-Squared Error
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
RSE	Relative Signal Enhancement
SD	Standard Deviation
SE	Signal Enhancement
SPGR	Spoiled Gradient Echo
SSIM	Structural Similarity Index
TCIA	The Cancer Imaging Archive
TM	Tofts Model
ULR	Univariate Logistic Regression

List of Publications

Original Publications

The following two original publications are subject of this cumulative dissertation in accordance with the promotion regulation for natural sciences of the medical faculty of the Ludwig-Maximilians-University Munich.

[1] **A. Mittermeier**, B. Ertl-Wagner, J. Ricke, O. Dietrich, and M. Ingrisch, "Bayesian pharmacokinetic modeling of dynamic contrast-enhanced magnetic resonance imaging: Validation and application," *Physics in Medicine & Biology*, vol. 64, no. 18, Sep. 2019, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab3a5a.

[2] **A. Mittermeier**, P. Reidler, M. P. Fabritius, B. Schachtner, P. Wesp, B. Ertl-Wagner, O. Dietrich, J. Ricke, L. Kellert, S. Tiedt, W. G. Kunz, and M. Ingrisch, "End-to-End Deep Learning Approach for Perfusion Data: A Proof-of-Concept Study to Classify Core Volume in Stroke CT," *Diagnostics*, vol. 12, no. 5, p. 1142, May 2022, ISSN: 2075-4418. DOI: 10.3390/diagnostics12051142.

Additional Original Publications

[1] N. Hainc, C. Federau, A. Tyndall, **A. Mittermeier**, A. Bink, C. Stippich, and T. Schubert, "Evaluation of the clinical utility of maximum intensity projections of 3D contrast-enhanced, T1-weighted imaging for the detection of brain metastases," *Cancer Reports (Hoboken, N.J.)*, vol. 3, no. 5, e1277, Oct. 2020, ISSN: 2573-8348. DOI: 10.1002/cnr2.1277.

[2] M. P. Fabritius, M. Seidensticker, J. Rueckel, C. Heinze, M. Pech, K. J. Paprottka, P. M. Paprottka, J. Topalis, A. Bender, J. Ricke, **A. Mittermeier**, and M. Ingrisch, "Bi-Centric Independent Validation of Outcome Prediction after Radioembolization of Primary and Secondary Liver Cancer," *Journal of Clinical Medicine*, vol. 10, no. 16, p. 3668, Jan. 2021. DOI: 10.3390/jcm10163668.

Conference Proceedings

[1] **A. Mittermeier**, C. Pirkl, M. J. Schneider, K. Parodi, and M. Ingrisch, "Pixel-wise evaluation of Intravoxel Incoherent Motion MRI: A Bayesian Hierarchical modeling approach," in *ESMRMB Congress 2017 – 34th Annual Scientific Meeting*, Barcelona, Spain, Oct. 2017.

[2] **A. Mittermeier**, O. Dietrich, B. Ertl-Wagner, and M. Ingrisch, "Bayes'sche pharmakokinetische Modellierung von DCE-MRT-Daten.," in *49. Jahrestagung der DGMP und 21. Jahrestagung der DS-ISMRM in Nürnberg*, Hrgs. M. Wucherer, C. Bert, F. Laun, 2018, p. 15, ISBN: 978-3-9816002-5-4.

[3] **A. Mittermeier**, B. Ertl-Wagner, O. Dietrich, and M. Ingrisch, "Bayesian Pharmacokinetic Modeling of Dynamic Contrast-Enhanced Magnetic Resonance Imaging: Validation and Application.," in *Proceedings of the Joint Annual Meeting ISMRM-ESMRMB (27th Annual Meeting & Exhibition)*, Montreal, QC, Canada, May 2019, p. 2353.

Abstract

Dynamic contrast-enhanced (DCE) imaging describes an imaging procedure in which a contrast agent (CA) is administered intravenously and several medical images are acquired over a period of time to show the blood flow (perfusion) through tissue. DCE magnetic resonance imaging (MRI) and computed tomography (CT) are frequently used in clinical routine, e.g., for stroke diagnosis, in oncology or cardiology. The acquired DCE-MRI and DCE-CT images are evaluated to derive hemodynamic parameters of the tissue of interest, e.g., blood flow or blood volume. The assessment of the hemodynamic state of a tissue can then help to, e.g., identify impaired perfusion or monitor therapy response in tumor therapy.

The basic principles for evaluating DCE imaging data are the same for MRI and CT and involve two independent steps. First, the time-dependent signal intensities are measured in a supplying artery as well as in the tissue of interest and are converted to CA concentrations. Then, the concentration-time curves are analyzed by an appropriate tracer-kinetic model which yields physiological, hemodynamic parameters of the tissue of interest. Reliable quantification of these hemodynamic parameters is essential, as estimation errors can have implications on clinical decision making. However, the evaluation of DCE imaging data is complex and depends on approximations that may not always apply, such as the correct tracer-kinetic model. This work aims to make perfusion quantification more robust by

(i) increasing the reliability of perfusion parameter estimates by incorporating the uncertainty of estimates using Bayesian modeling and

(ii) skipping tracer-kinetic modeling completely to directly predict clinical endpoints using a deep-learning (DL) approach.

In a first proof-of-concept study, Bayesian tracer-kinetic modeling of DCE-MRI data was investigated with the purpose to validate accuracy and precision of the parameter estimates, as well as to assess therapy response for breast cancer patients from parameter estimates. To this end, the tracer-kinetic Tofts model was implemented in a Bayesian probabilistic framework to calculate probability distributions of the perfusion parameters and determine their uncertainty. This Bayesian Tofts model

(BTM) allows to incorporate domain knowledge about hemodynamics by means of a *prior* probability distribution and combines it with measured data to derive an information-enriched *posterior* parameter distribution.

The estimates of the BTM were validated against conventional non-linear least squares (NLLS) tracer-kinetic analysis with additional bootstrapping as a measure of uncertainty in a simulation setting using a digital DCE-MRI reference object with known ground truth. The accuracy of the estimated parameter map for the exchange rate K^{trans} was assessed by means of the structural similarity index (SSIM) which showed high accordance with the ground truth for both approaches (SSIM_{BTM} = 96%, $\text{SSIM}_{\text{NLLS}}$ = 91%). The precision of the BTM posterior distributions was found to be in good agreement with the bootstrapped NLLS approach (SSIM = 91%), which verifies that the posterior distribution of the BTM yields sensible uncertainty estimates.

The BTM was then applied to a breast cancer DCE-MRI dataset to assess the treatment response during preoperative neoadjuvant chemotherapy (NACT) and to discriminate between responder and non-responder. To this end, K^{trans} posterior probability distributions were compared across visits using Cohen's d, which incorporates the standard deviation of the posterior. Assessment of treatment response was found to be excellent by means of a receiver operating characteristics (ROC) area under the curve (AUC) of 0.952. In summary, the Bayesian modeling approach provides robust, intrinsic uncertainty estimates in form of a posterior distribution and yields precise information about changes in parameter estimates.

To address the problem of tracer-kinetic modeling assumptions, we investigated the direct prediction of clinical endpoints from raw DCE-CT data using an end-to-end DL approach in a second proof-of-concept study. The aim of this study was to develop and train a dedicated convolutional neural network (CNN) that can learn hemodynamic information from time-resolved CT acquisitions to completely circumvent conventional tracer-kinetic analysis. Specifically, we aimed to discriminate a cohort of acute ischemic stroke patients by small or large core volume in a simplified classification task without tracer-kinetic modeling.

The CNN model architecture comprises two identical submodels to extract spatial and temporal features from two 2D+t CT slices which are combined to predict the dichotomized core volume. Model training was performed in a 10-fold cross validation setting and evaluation on the test set yielded a mean (standard deviation) ROC AUC of 0.72(0.10). A final model was created using an ensemble approach and was independently validated on the external ISLES 2018 challenge dataset (ROC AUC = 0.61). In summary, the proposed end-to-end DL approach circumvents conventional tracer-kinetic modeling and allows the prediction of dichotomized infarct core volume from DCE-CT images without underlying tracer kinetic assumptions. The DL model developed in this proof-of-concept study lays the groundwork for further adaptation to additional clinically relevant endpoints.

In conclusion, this works demonstrates two novel approaches for a robust evaluation of DCE imaging data for perfusion quantification. On the one hand, the confidence in perfusion parameter estimates is increased by considering the intrinsic uncertainty of estimates introduced by a Bayesian formulation of tracer-kinetic modeling. On the other hand, we showed that clinically relevant endpoints can be learned directly from raw DCE imaging data, circumventing tracer-kinetic analysis, through the use of specialized convolutional neural networks. The combination of both approaches in so-called Bayesian neural networks is an emerging research area and could be suitable to increase the robustness of perfusion quantification and therefore clinical acceptance even further.

Zusammenfassung

Die dynamische kontrastverstärkte Bildgebung (*dynamic contrast-enhanced*, DCE) ist ein bildgebendes Verfahren, bei dem ein Kontrastmittel (*contrast agent*, CA) intravenös verabreicht wird und Bilddaten über einen bestimmten Zeitraum aufgenommen werden, um den Blutfluss (Perfusion) durch das Gewebe darzustellen. DCE-Magnetresonanztomographie (MRT) und DCE-Computertomographie (CT) werden häufig in der klinischen Routine eingesetzt, z. B. zur Schlaganfalldiagnose, in der Onkologie oder Kardiologie. Die aufgenommenen DCE-MRT- und DCE-CT-Bilder müssen ausgewertet werden, um hämodynamische Parameter des Gewebes abzuleiten, z. B. den Blutfluss oder das Blutvolumen. Die Beurteilung des hämodynamischen Zustands eines Gewebes kann dann helfen, Durchblutungsstörungen zu erkennen oder das Therapieansprechen bei einer Tumortherapie zu überwachen.

Die Grundprinzipien für die Auswertung von DCE-Bilddaten sind für MRT und CT gleich und umfassen zwei unabhängige Schritte. Zunächst werden die zeitabhängigen Signalintensitäten sowohl in einer versorgenden Arterie als auch im Gewebe gemessen und in CA-Konzentrationen umgerechnet. Anschließend werden die Konzentrations-Zeit-Kurven mit Hilfe eines geeigneten Tracer-Kinetik-Modells analysiert, um physiologische, hämodynamische Parameter des Gewebes zu ermitteln. Eine zuverlässige Quantifizierung dieser hämodynamischen Parameter ist unerlässlich, da Schätzfehler Auswirkungen auf die klinische Entscheidungsfindung haben können. Die Auswertung von DCE-Bilddaten ist komplex und beruht auf Annäherungen, die nicht immer zutreffen, wie beispielsweise die richtige Wahl des Tracer-Kinetik-Modells. Diese Arbeit hat das Ziel, die Perfusionsquantifizierung robuster zu gestalten, indem

(i) die Zuverlässigkeit der Parameterschätzungen durch miteinbeziehen der Unsicherheit der Schätzungen mit Hilfe von Bayes'scher Modellierung erhöht wird und

(ii) die Tracer-Kinetik-Modellierung vollständig übersprungen wird, um klinische Endpunkte mit Hilfe eines Deep-Learning-(DL-)Ansatzes direkt vorherzusagen.

In einer ersten Proof-of-Concept-Studie wurde die Bayes'sche tracer-kinetische Modellierung von DCE-MRT-Daten mit dem Ziel untersucht, die Genauigkeit und Präzi-

sion der Parameterschätzungen zu validieren und das Therapieansprechen bei Brustkrebspatientinnen mit Hilfe der geschätzen Parameter zu bewerten. Zu diesem Zweck wurde das tracer-kinetische Tofts-Modell in einer Bayes'schen probabilistischen Umgebung implementiert, um die Unsicherheit der Perfusionsparameterschätzungen zu berechnen. Dieses Bayes'sche Tofts-Modell (BTM) ermöglicht es, Domänenwissen über die Hämodynamik des Gewebes mittels einer *A-priori*-Wahrscheinlichkeitsverteilung, kurz Prior, einzubeziehen. Kombiniert mit den gemessenen Daten, kann eine spezifischere *A-posteriori*-Verteilung der Parameter, kurz Posterior, abgeleitet werden.

Die Schätzungen des BTM wurden gegenüber der konventionellen Tracer-Kinetik-Analyse (*non-linear least squares*, NLLS) mit zusätzlichem Bootstrapping als Maß für die Unsicherheit validiert. In einer Simulationsstudie wurde dafür ein digitales DCE-MRT-Referenzobjekt mit bekannter *ground truth* verwendet. Die Genauigkeit der geschätzten Parameterkarte für die Austauschrate K^{trans} wurde mit Hilfe des *structural similarity index*(SSIM) bewertet, der für beide Ansätze eine hohe Übereinstimmung mit der *ground truth* zeigte ($\text{SSIM}_{\text{BTM}} = 96\%$, $\text{SSIM}_{\text{NLLS}} = 91\%$). Die Präzision der Posterior-Verteilung stimmte gut mit dem NLLS-Bootstrapping-Ansatz überein (SSIM = 91%), d.h. die Posterior-Verteilung des BTM liefert sinnvolle Unsicherheitsabschätzungen.

Das BTM wurde dann auf einen Brustkrebs-DCE-MRT-Datensatz angewandt, um das Ansprechen auf die Behandlung während einer präoperativen neoadjuvanten Chemotherapie (NACT) zu bewerten und zwischen *respondern* und *non-respondern* zu unterscheiden. Zu diesem Zweck wurden die K^{trans}-Posterior-Wahrscheinlichkeitsverteilungen über die verschiedenen Besuche hinweg unter Verwendung von Cohens d verglichen, wodurch die Standardabweichung des Posteriors mit einfließt. Die Bewertung des Therapieansprechens wurde mit einer Fläche unter der *receiver operating characteristics* Kurve (ROC AUC) von 0,952 als hervorragend eingestuft. Zusammenfassend liefert der Bayes'sche Modellierungsansatz robuste, intrinsische Unsicherheitsschätzungen in Form einer Posterior-Verteilung und präzise Informationen über Änderungen der Parameterschätzungen.

Um das Approximationsproblem bei der tracer-kinetischen Modellierung zu adressieren, wurde in einer zweiten Proof-of-Concept-Studie die direkte Vorhersage von klinischen Endpunkten aus DCE-CT-Rohdaten unter Verwendung eines End-to-End-DL-Ansatzes untersucht. Ziel dieser Studie war es, ein spezielles neuronales Netzwerk (*convolutional neural network*, CNN) zu entwickeln und zu trainieren, das hämodynamische Informationen aus zeitaufgelösten CT-Aufnahmen lernen kann, um die konventionelle tracer-kinetische Analyse vollständig zu umgehen. Konkret sollte eine Kohorte von Patienten mit akutem ischämischem Schlaganfall in einer vereinfachten Klassifizierungsaufgabe ohne tracer-kinetische Modellierung nach kleinem oder großem Infarktkernvolumen unterschieden werden.

Die CNN-Modellarchitektur besteht aus zwei identischen Teilmodellen zur Extrak-

tion räumlicher und zeitlicher Merkmale aus zwei 2D+t CT-Schichten, die zur Vorhersage des dichotomen Infarktkernvolumens kombiniert werden. Das Training des Modells wurde in einer 10-fachen Kreuzvalidierung durchgeführt, und die Auswertung des Testdatensatzes ergab eine mittlere (Standardabweichung) ROC AUC von 0,72(0,10). Ein finales Modell wurde mithilfe eines Ensemble-Ansatzes erstellt und unabhängig auf dem externen ISLES 2018 Challenge-Datensatz validiert (ROC AUC = 0,61). Zusammenfassend umgeht der vorgeschlagene End-to-End-DL-Ansatz die konventionelle Perfusionsanalyse und ermöglicht die Vorhersage des dichotomen Infarktkernvolumens aus DCE-CT-Bildern ohne zugrunde liegende Tracer-Kinetik-Annahmen. Das in dieser Proof-of-Concept-Studie entwickelte DL-Modell legt den Grundstein für weitere Anpassungen an zusätzliche klinisch relevante Endpunkte.

Zusammengefasst konnte in dieser Arbeit eine robuste Auswertung von DCE-Bilddaten für die Perfusionsquantifizierung mit Hilfe zweier neuartiger Analysemethoden demonstriert werden. Einerseits wird das Vertrauen in die Perfusionsparameterschätzungen erhöht, indem die intrinsische Unsicherheit der Schätzungen berücksichtigt wird, die durch eine Bayes'sche Formulierung der Tracer-Kinetik-Modellierung ermittelt werden kann. Andererseits konnten wir zeigen, dass klinische Endpunkte direkt aus den Rohdaten der DCE-Bildgebung durch den Einsatz von neuronalen Netzen gelernt werden können, um so die Tracer-Kinetik-Analyse komplett zu umgehen. Die Kombination beider Ansätze in sogenannten Bayes'schen neuronalen Netzen ist eine aufstrebende Forschungsrichtung und könnte geeignet sein, die Robustheit der Perfusionsquantifizierung und dadurch die klinische Akzeptanz noch weiter zu erhöhen.

1 Introduction

In the context of medical imaging, perfusion is defined as the delivery of blood to the capillary beds and is measured as the rate of delivery per unit of time and tissue [1]. In contrast to the bulk blood flow in arteries and veins, perfusion describes blood flow at the capillary level which controls the body's metabolism by supplying tissue and cells with nutrients and removing waste products. Thus, it is an important parameter to characterize the hemodynamic status of the microvasculature. Further hemodynamic parameters include the density of blood vessels, the permeability of the vessel walls or the mean transit time of the blood.

Impaired or disrupted perfusion results in a lack of blood supply and is associated with various pathologies. Hypoperfused brain tissue for example is related to ischemic stroke [2] and malperfused cardiovascular tissue can lead to chronic heart diseases such as myocardial infarction [3]. Furthermore, the assessment of the hemodynamic state of tumors is of great interest to understand growth and proliferation compared to normal tissue or monitor therapy response [4–6].

The assessment of the hemodynamics of tissue can be differentiated into contrast agent (CA) free methods on the one hand, e.g., diffusion weighted imaging [7] or arterial spin labeling in magnetic resonance imaging (MRI) [8] and dynamic contrast-enhanced (DCE) imaging approaches [9] on the other hand, available in, e.g., MRI and computed tomography (CT). Due to the administration of CA, the latter are so-called bolus-tracking experiments which have their origin in nuclear medicine and, for medical imaging, date back to the 1980s [10, 11]. In this work, the focus lies on the assessment of hemodynamics using DCE imaging and the subsequent quantification of tissue perfusion.

To quantify perfusion in a bolus-tracking experiment, CA is first injected intravenously and multiple 2D or 3D images are acquired over a certain time range. These time-resolved 2D+t or 3D+t images show the passage of contrast agent through tissue, indicated by temporally varying signal intensities in the proximity of the tracer. Next, the time-dependent signal intensities are measured in a supplying artery and in the tissue of interest and are converted to CA concentrations. The concentration-time curves are then analyzed by an appropriate tracer-kinetic model which yields

physiological hemodynamic parameters of the tissue, e.g. relative blood volume or blood flow [12–16].

Reliable quantification of hemodynamic parameters is essential, as estimation errors can have implications on clinical decision making [17]. However, the evaluation of contrast-enhanced imaging is complex and depends on approximations that may not always apply, e.g., the right localization of the supplying artery, the estimation of CA concentration or the correct choice of a tracer-kinetic model. This work aims to make perfusion quantification more robust by i) calculating uncertainty in perfusion parameter estimates using Bayesian modeling and ii) skipping tracer-kinetic modeling completely to directly predict clinical endpoints using a deep learning approach.

The remainder of this work is organized as follows: chapter 2 provides the necessary background to understand perfusion imaging and tracer-kinetic analysis for MRI and CT. It elaborates further on DCE analysis methods including basics of Bayesian statistics and deep learning for image classification. In chapter 3, the author's contributions to the original publications (chapter 4 and chapter 5) can be found.

2 | Background

2.1 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) of biological tissue is a non-invasive imaging modality based on nuclear magnetic resonance (NMR). The aim is to reconstruct an image from the spatial distribution and relaxation characteristics of, most frequently, hydrogen nuclei ^{1}H. This section provides an overview about the fundamentals of MRI and is based on standard textbooks [18, 19].

2.1.1 Nuclear Magnetic Resonance

The principle of NMR is the interaction of nuclear spin, denoted as the angular momentum quantum number I, with an external magnetic field. The magnetic moment of a nucleus $\boldsymbol{\mu}$ is connected to a non-zero nuclear spin $\boldsymbol{I}$ by the nucleus-specific *gyromagnetic ratio* γ via

$$\boldsymbol{\mu} = \gamma \boldsymbol{I}. \tag{2.1}$$

When exposed to an external magnetic field $\boldsymbol{B}$, the behavior of a nucleus can be described by the classical equation for a magnetic dipole

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \gamma \boldsymbol{\mu} \times \boldsymbol{B}. \tag{2.2}$$

It follows that the hydrogen nuclei start to precess about the direction of the magnetic field with the *Larmor frequency*

$$\omega_0 = \gamma B. \tag{2.3}$$

When an ensemble of nuclei is exposed to a static magnetic field $\boldsymbol{B_0} = B_0 \hat{\boldsymbol{e}}_{\boldsymbol{z}}$, a macroscopic magnetization gradually builds up. As illustrated in Fig. 2.1, the isotropic distribution of the ensemble of spins in a liquid sample is slightly skewed due to the magnetic field, introducing a *net magnetization* $\boldsymbol{M}(t)$. The dynamic properties of the macroscopic magnetization can be described by the same equations of motion as for a single nucleus as

$$\frac{\mathrm{d}\boldsymbol{M}(t)}{\mathrm{d}t} = \gamma \boldsymbol{M}(t) \times \boldsymbol{B}(t). \tag{2.4}$$

Particularly, the transverse part of the magnetization also precesses with the Larmor frequency ω_0 about the direction of the magnetic field.

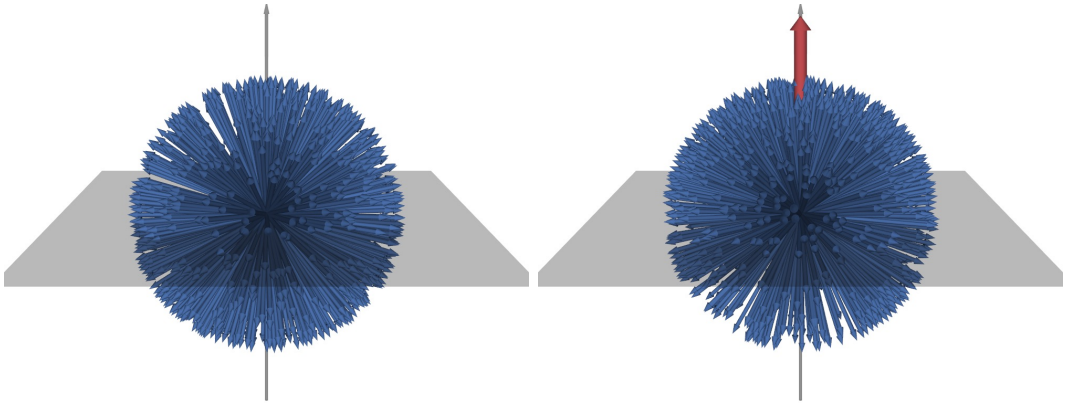

Figure 2.1: Angular spin distribution for an ensemble of spins in a liquid sample. The distribution is isotropic without an external magnetic field (left) and gets slightly skewed when exposed to a static magnetic field $\boldsymbol{B_0}$ (right), yielding a net magnetization $\boldsymbol{M}(t)$ (red arrow).

2.1.2 Excitation and Relaxation

To measure a signal in MRI, the macroscopic magnetization needs to be excited by an additional oscillating magnetic field $\boldsymbol{B_1}(t)$. The oscillating field $\boldsymbol{B_1}(t)$ is set to rotate about the direction $\hat{\boldsymbol{e}}_{\boldsymbol{z}}$ of the static magnetic field, tuned to the Larmor frequency ω_0. Applying $\boldsymbol{B_1}(t)$ for a certain amplitude and time Δt, defines a radiofrequency (RF) pulse which perturbs the macroscopic magnetization from the z-axis by the *flip angle*

$$\alpha = \gamma B_1 \Delta t. \tag{2.5}$$

After excitation with a 90° RF pulse, the magnetization $\boldsymbol{M}$ precesses in the transverse xy-plane and generates a RF field itself. This electromagnetic field induces an electric current in the coils which constitutes the signal in MRI. Immediately after excitation, the magnetization starts to relax until it returns to its equilibrium magnetization $\boldsymbol{M}_0 = M_0 \hat{\boldsymbol{e}}_{\boldsymbol{z}}$. This *relaxation* process is caused by interactions of the excited spins with their surrounding (spin-lattice) and with each other (spin-spin).

The spin-lattice or *longitudinal* relaxation time T_1 describes the exponential recovery of the longitudinal part of the magnetization M_z. The energy gained during the relaxation process is transferred as heat to the lattice. The spin-spin or *transverse* relaxation time T_2 characterizes the irreversible exponential decay of the transverse part of the magnetization M_{xy} caused by a loss of phase coherence between the spins. The actual decay is faster than explained by T_2 alone because of additional magnetic field inhomogeneities. This reversible part of the relaxation is denoted as

T_2'-relaxation and the entire decay of M_{xy}, referred to as *free induction decay* (FID), is characterized by the time constant T_2^*, defined as

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'}. \tag{2.6}$$

2.1.3 Echo Formation

As introduced above, the T_2'-relaxation caused by static local inhomogeneities of the external magnetic field can be reversed. Depending on their spatial position, the spins rotate at slightly different frequencies and start to dephase. By applying an 180°-RF pulse, the magnetization vectors and phases of the spins are inverted. The same static inhomogeneities now cause the spins to rephase, forming a *spin echo* at the echo time T_E.

The generation of gradient echoes is governed by similar principles of dephasing and rephasing, but relies on the application of an additional magnetic field gradient G, superimposed with the static magnetic field B_0. As illustrated in Fig. 2.2, a constant negative gradient $-G$ along the z-axis in the time interval (t_1, t_2) causes a phase accumulation for spins at position z. This phase accumulation is compensated by the inverse gradient $+G$, applied during the time interval (t_3, t_4), forming a *gradient echo* at timepoint T_E.

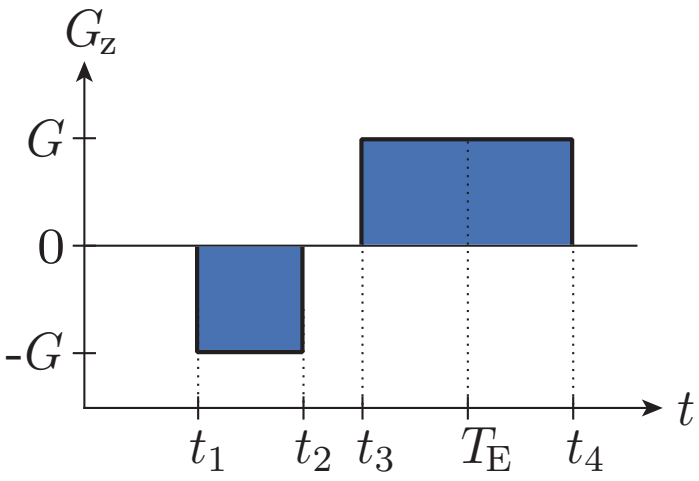

Figure 2.2: Scheme for a gradient echo in a 1D MRI experiment. The gradient echo forms at the time point TE, exactly when the integral of the first negative gradient lobe cancels the second positive gradient lobe. Adapted from [20].

2.1.4 Spatial Encoding and Image Reconstruction

Spins that are exposed to the same static magnetic field B_0 exhibit the same Larmor frequency ω_0. In order to get information about the location of a spin in a sample, it is possible to spatially encode the spins with an additional gradient field; the external magnetic field becomes then

$$\boldsymbol{B}(\boldsymbol{x}, t) = \boldsymbol{B_0} + \boldsymbol{G}(t) \cdot \boldsymbol{x}\hat{\boldsymbol{e}}_{\boldsymbol{z}}. \tag{2.7}$$

Assuming a gradient field $G_x(t)$ with spatial variation in x, then

$$\omega_G(x, t) = \gamma x G_x(t) \tag{2.8}$$

defines a *frequency encoding* of spins along x and links position to precession rate. Likewise, the transverse part of the magnetization $M_{xy}(x, t)$ now depends on time and location. The detected signal $S(t)$ is proportional to the transverse magnetization integrated along x,

$$S(t) \propto \int \mathrm{d}x M_{xy}(x, t), \tag{2.9}$$

which takes on the form of a Fourier transform of the spin density $\rho(x)$ at the point $k(t)$. Since $k(t)$ is determined by $G(t)$, tuning the gradient field allows to measure $S(t) = S(k(t))$ (neglecting a scaling factor) at arbitrary points k_i.

This principle can be generalized to two dimensions by introducing another gradient field $G_y(t)$ along the y axis which defines the so-called *phase encoding*. The signal can then be measured in an equidistantly spaced grid of $N_x \cdot N_y$ points, illustrated in Fig. 2.3. The distance between two points Δk determines the field of view (FOV) and since the discrete Fourier transform is reversible, also determines the spatial resolution of the reconstructed MR image as

$$\Delta x = \frac{FOV}{N} = \frac{2\pi}{N\Delta k}. \tag{2.10}$$

In three dimensions, $k(t)$ is extended to a three dimensional vector $\boldsymbol{k}(t)$. The resulting vector space is called *k-space*. Image reconstruction can be performed by inverse (discrete) fast Fourier transform of the acquired k-space data.

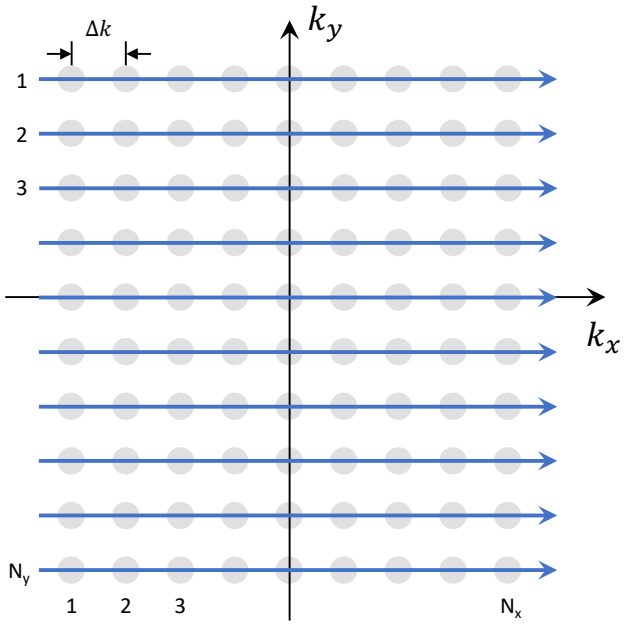

Figure 2.3: Cartesian k-space sampling in a 2D slice. In total, $N_x \cdot N_y$ points are sampled, the distance between two points is denoted as Δk.

2.2 Computed Tomography Imaging

X-ray computed tomography (CT) is a medical imaging technique which uses a rotating X-ray tube and detectors to generate digital, cross-sectional images. The basic principle of CT imaging is the measurement of X-ray attenuation of a physical object from different directions to compute an image. This section provides an overview about the fundamentals of CT imaging and is based on standard textbooks [21, 22].

2.2.1 X-ray Generation

The X-rays for CT imaging originate from a cathode tube with applied voltages between 25 and 150 keV. The temperature of the source metal must be high enough to overcome the binding energy of the electrons, creating an electron cloud at the filament. The tube voltage accelerates the electrons to the anode where they are abruptly stopped by interactions with the orbital electrons and the atomic nuclei of the anode material.

The electron-matter interaction at the anode creates *bremsstrahlung*, a multi-process cascade which represents the main contribution to the X-ray spectrum, illustrated in Fig. 2.4. Superimposed is a characteristic line spectrum which originates from ionization and subsequent gamma emission of the anode material. Direct electron-nucleus

interaction is relatively rare and constitutes to the upper limit of the spectrum.

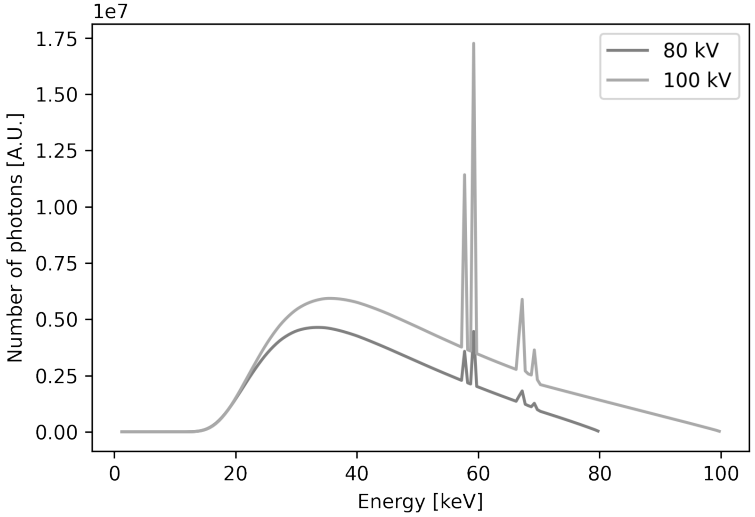

Figure 2.4: X-ray spectrum of a tungsten anode at acceleration voltages of $U_a = 80$ and 100kV. The anode angle is 10° and 2mm Al filtering has been applied. The plots shows the continuous *bremsstrahlung* and the superimposed characteristic peaks. Created using SpekPy version 2.0.7 [23].

The generation of X-rays by electron deceleration is a rare process. Within a conventional anode, approximately 99% of the kinetic energy of the electrons is converted to heat. Furthermore, the intensity of the *bremsstrahlung* is proportional to the atomic number Z of the anode material. Tungsten fulfills the requirements for high heat load and large Z and is therefore often used as anode material.

The quality of the X-ray beam, i.e., size and shape of the X-ray focus, determines the quality of the generated image. Relativistic effects of high energy electrons and self-absorption of the generated photons on the anode surface can lead to intensity reduction at the detector, deteriorating the image. Furthermore, the image quality is influenced by the attenuation behavior of the X-rays traveling through matter. Lower-energy X-rays of the poly-chromatic spectrum are attenuated more strongly which results in the so called *beam-hardening* effect. Applying a thin metal layer beam filter on the anode results in decreased intensity but increased average energy which mitigates downstream beam-hardening artifacts.

2.2.2 Photon–Matter Interaction

The interaction of X-rays with matter causes attenuation of the initial photon intensity. The following principal, competing physical processes contribute to this

macroscopic attenuation: Rayleigh scattering, photoelectric absorption, Compton scattering and pair production.

Rayleigh scattering is the elastic scattering of photons if the diameter of the nucleus is small compared to the wavelength of the incident photon. This process is observed mainly at low photon energies. While only the direction of the scattered photon changes, no energy is transferred.

Photoelectric absorption describes the total absorption of photon energy $h\nu$ by a bound electron. The interacting electron is emitted from the atom with the incident photon energy minus the binding energy of the electron in an ionization process. The vacant space is filled by a nearby free electron and as a result, characteristic X-ray fluorescence lines can be measured. If the fluorescence X-ray radiation is absorbed by another electron, the so-called Auger process is triggered. The photoelectric effect cross-section has a strong dependence on the photon energy $h\nu$ and the atomic number Z,

$$\sigma_{\text{Photo}} \propto \frac{Z^4}{h\nu^3} \tag{2.11}$$

(for photon energies below 50 keV), which is the reason for the choice of high-Z elements as contrast media, e.g., iodine (I, Z=53).

Compton scattering is the inelastic scattering of photons on quasi-free electrons. Energy of the incident photon is transferred to the scattered electron which is emitted from the atom. The scattered photon has a reduced wavelength $\Delta\lambda$ which depends on the scatter angle θ. The Compton scattering cross section depends linearly on the atomic number

$$\sigma_{\text{Compton}} \propto Z. \tag{2.12}$$

Pair production is the dominant process for high photon energies $h\nu > 2\times$ 511 keV. After a short distance, the positron annihilates with an electron which produces two γ-rays traveling in approximately opposite directions with 511 keV each.

The linear attenuation coefficient μ is composed by the individual cross sections and can be calculated as

$$\mu = \frac{\rho N_{\text{A}}}{A} \cdot \sigma_{\text{tot}}, \tag{2.13}$$

with density ρ, atomic weight A, Avogadro constant N_{A} and the total cross section σ_{tot}. The total attenuation of X-rays in matter depends on the penetrated material and the energy of the photons. The energy of photons in diagnostic imaging ranges between 20 and 150 keV, therefore the photoelectric effect and Compton scattering are the dominant contributions to the total attenuation. Fig. 2.5 shows the individual contributions for water in the diagnostic energy window of CT.

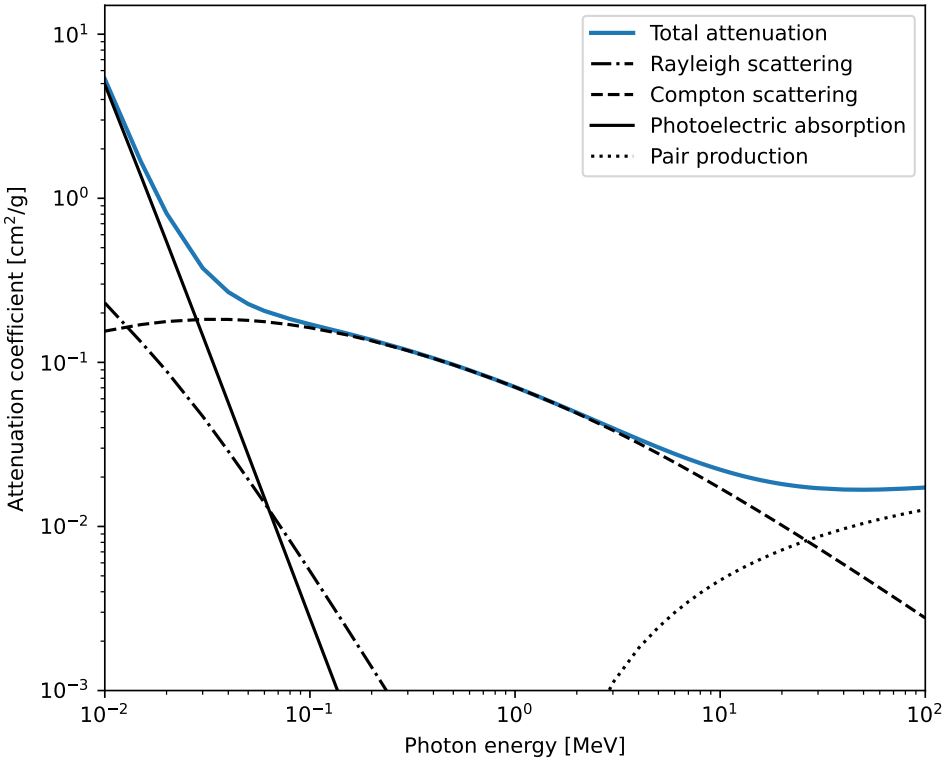

Figure 2.5: Mass attenuation coefficient μ/ρ over incident photon energy $h\nu$ for H_2O in the energy range of diagnostic imaging. Data source: Photon Cross Sections Database (XCOM) [24].

2.2.3 Lambert-Beer's Law and X-ray detection

Assuming a monochromatic X-ray beam penetrating a homogeneous object with thickness s, the radiation intensity can be calculated by Lambert-Beer's law of attenuation:

$$I(s) = I_0 e^{-\mu s} \tag{2.14}$$

with the single, constant attenuation coefficient μ and the measured intensity $I(s)$.

Problems with Lambert–Beer's law arise when dealing with non-homogeneous material, as illustrated in Fig. 2.6. If the attenuation coefficient is dependent on the position within the penetrated material $\mu(\eta)$, then equation 2.14 is no longer correct. Instead, accounting for spatially varying attenuation over a distance s along the path

of the X-ray beam η, the intensity calculates as

$$I(s) = I(0)\mathrm{e}^{-\int_0^s \mu(\eta)\mathrm{d}\eta}. \tag{2.15}$$

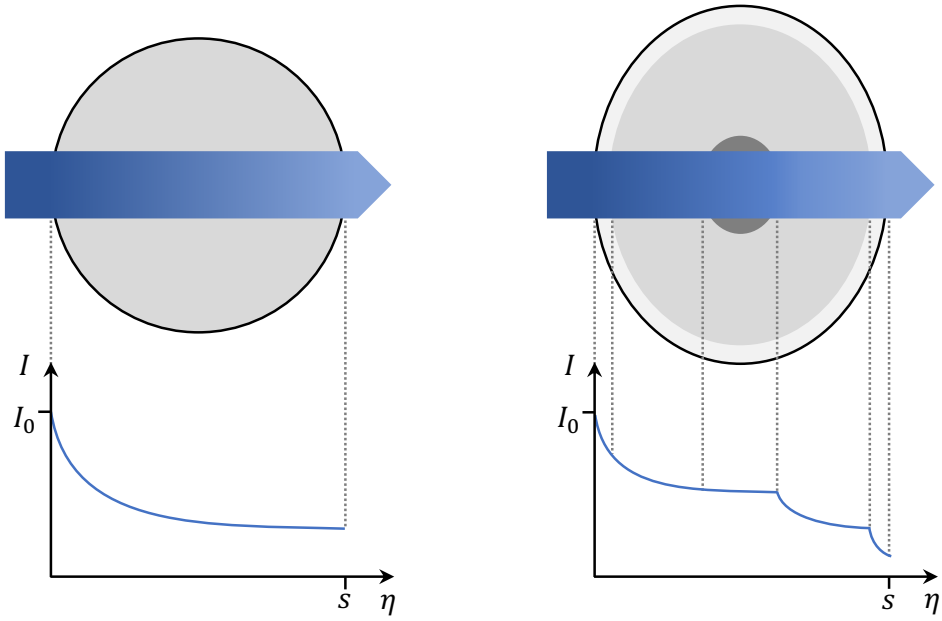

Figure 2.6: Attenuation of radiation intensity I of monochromatic X-rays passing through matter with thickness s along the path η. The left side corresponds to equation 2.14 for a homogeneous object, the right side represents equation 2.15 for inhomogeneous objects. Adapted from [21].

X-rays are detected indirectly via interaction with the detector material after traveling through matter. This interaction follows the same principles of photon-matter interaction described above. The quality of the detector depends on the geometric and quantum efficiency, in other words, the conversion rate of X-rays that contribute to the signal. Modern CTs are equipped with solid-state scintillator detectors which have a high quantum efficiency, i.e., a high energy conversion due to photoelectric absorption. However, a relatively thick collimator grid is needed to prevent the detection of scattered X-rays, which in turn leads to a reduced spatial detector resolution. Solid-state flat-panel detectors are the extension of scintillator detectors to multiple rows or slices. This increases the geometric efficiency as cylindrical detector systems are adapted to the cone-beam source.

2.2.4 Image Acquisition

A CT image of an object is acquired by measuring a finite number of line integrals along the path of an X-ray beam η to the detector to determine the distribution of the attenuation coefficient $\mu(\xi, \eta)$ within the object (Fig. 2.7). Although modern

CT scanners acquire data in a cone-beam geometry or spiral, the basics of image reconstruction are best explained for parallel beam geometry. Assuming a fixed projection angle ϕ, the X-ray source is shifted along the axis ξ and the X-rays traveling through the object along η create an attenuation profile. For a particular projection angle ϕ and a linear source position ξ, the projection integral is given as

$$p_\phi(\xi) = \int_0^s \mu(\xi, \eta) \mathrm{d}\eta \tag{2.16}$$

Subsequently, the X-ray source and detector unit is rotated further and shifted again, until the object to be reconstructed is illuminated over a range of at least 180°. The full set of projections is called the Radon transform of the image or sinogram.

The aim is to reconstruct the object function $f(x, y)$ from the measured projection data $p_\phi(\xi)$. Therefore, the projections are transformed into frequency space (u, v) using a 1D Fourier Transform $\mathcal{F}_{1d}$. The Fourier slice theorem allows us to identify each projection profile with a radial line in the Cartesian frequency space and superimpose them according to their projection angle ϕ. Now, an inverse 2D Fourier transform $\mathcal{F}_{2d}^{-1}$ is carried out to reconstruct the object function $f(x, y)$ which yields the distribution of attenuation values as a function of the fixed patient coordinate system (x, y). This simple backprojection introduces blurring of the reconstructed image as each point in the projection contributes to the whole image instead of its original 2D position. To counteract this unsharpening, each projection is convolved with a high pass filter before backprojection. Next to this filtered backprojection reconstruction approach, most modern CT scanner use iterative reconstruction to obtain the object function from the projection data.

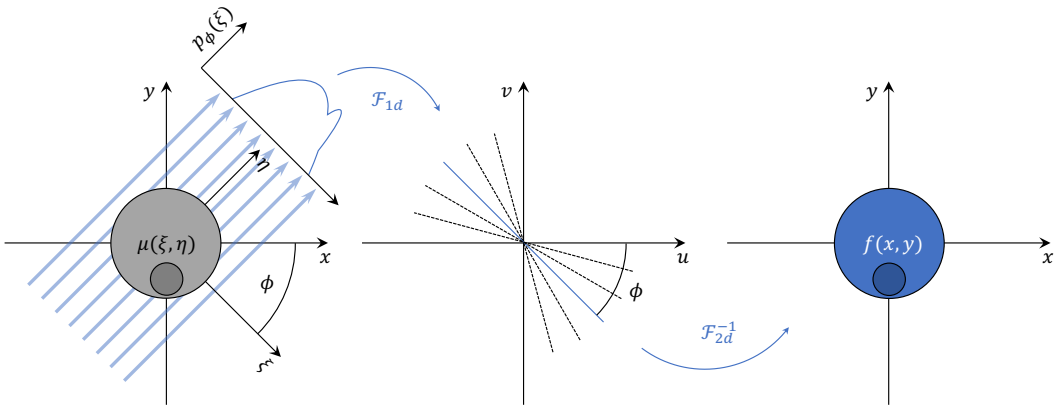

Figure 2.7: Schematic illustration for CT image reconstruction. An object with unknown attenuation coefficient $\mu(\xi, \eta)$ is radiated with X-rays from multiple projection angles ϕ over a range of at least 180°. For a fixed projection angle ϕ, the X-ray source is shifted along ξ and the X-rays traveling through the object along η create an attenuation profile - the projection $p_\phi(\xi)$. The projections are then transformed into frequency space (u, v) using a 1D Fourier Transform $\mathcal{F}_{1d}$ and superimposed according to their projection angle. Finally, an inverse 2D Fourier transform $\mathcal{F}_{2d}^{-1}$ is carried out to reconstruct the object function $f(x, y)$ which yields the distribution of attenuation values as a function of the fixed patient coordinate system.

2.3 Dynamic Contrast-Enhanced Imaging

Dynamic contrast-enhanced (DCE) imaging is an imaging technique based on the dynamic acquisition of multiple images in order to assess and quantify hemodynamic parameters of tissue perfusion. Perfusion describes the passage of fluid through the circulatory system to an organ or tissue. In particular, it refers to the delivery of blood to the capillary bed and is measured as volume per unit time per unit tissue mass. To make this passage visible, contrast agents (CA) are usually injected intravenously and images are acquired before, during and after the administration of CA [16]. The CA acts as an indicator and causes time-dependent changes in signal intensity. Analysis and quantification of the measured signals in DCE imaging relies on two independent steps:

1. Time-dependent signal intensities are measured in a supplying artery and the tissue of interest and are then converted to CA concentrations using dedicated signal theory of the respective imaging modality.

2. Time-concentration curves are analyzed using tracer-kinetic theory to derive hemodynamic parameters.

2.3.1 Data Acquisition

DCE data acquisition aims at providing time-resolved signal intensities $S(t)$ of the passage of CA through tissue of interest. For that purpose, multiple images need to be acquired for a specific period of time, balancing various competing requirements. Spatial coverage, for example, needs to be large enough to include the tissue regions of interest and its supplying artery, while the spatial resolution must be high enough to delineate relevant tissue structures. Likewise, acquisition time needs to be long enough to cover the slowest perfusion processes, while the temporal resolution, i.e., the time between two acquisitions, must be faster than the fastest processes [25]. At the same time, image noise needs to be minimized, e.g., by sacrificing spatial and temporal resolution, such that the contrast-to-noise ratio is sufficient to detect signal enhancement induced by the CA.

Additional difficulties arise, when measuring the CA concentration in the tissue-feeding artery, often referred to as arterial input function or AIF. To be able to accurately apply tracer-kinetic analysis (see subsection 2.3.3), the AIF needs to be measured precisely at the inlet to the tissue of interest. In practice, a single AIF is measured at a sufficiently large artery and bolus dispersion between tissue pixels and the site of measurement is often ignored, resulting in an error in hemodynamic parameter estimation [26]. In addition, measuring the AIF yields CA concentrations for the whole arterial blood. To obtain the concentration for arterial blood plasma, the scaling factor $(1 - hct)$ is applied to the measured concentration, with hct representing the hematocrit or volume percentage of red blood cells.

Furthermore, partial-volume effects are observed for arteries that are smaller than the order of magnitude of the spatial resolution [27]. Underestimations due to the influence of surrounding tissue can be corrected by a reference measurement in a larger vein. In particular, this applies for DCE imaging in the brain where the AIF is sometimes measured in the sagittal sinus, because feeding arteries are too small. Using an AIF from a draining vessel for analysis can be justified, as long as the time scale between two images is large enough so that differences between arterial and venous timing can be neglected. As a last resort, a population-averaged AIF can be used to approximate the CA concentration in the supplying artery [28].

In DCE-MRI, the most common CA is based on Gadolinium Gd^{3+} which has strong paramagnetic properties. Through dipole-dipole interactions with nearby water protons, the T_1 and T_2 relaxation times are reduced in proportion to CA concentration, thereby increasing signal intensity, depending on the pulse sequence. In DCE-CT, an iodine-based CA is used which increases X-ray attenuation linear to its concentration. The first step in DCE analysis is therefore to calculate the time-dependent CA concentration from the unique changes in signal intensity.

2.3.2 Converting signal to concentration

Signal Enhancement Assuming that the CA concentration $c(t)$ is proportional to the change in signal intensity, it can be approximated by the *signal enhancement* (SE) $S(t) - S_0$:

$$c(t) = k \cdot (S(t) - S_0). \tag{2.17}$$

Here, S_0 is the precontrast signal acquired before the administration of CA, often averaged over multiple precontrast timepoints and k is an a priori unknown constant. Assuming that k is tissue independent, i.e., that it is equal in the supplying artery and the tissue of interest, it cancels out when applying tracer-kinetic theory (see section 2.3.3). For DCE-CT, this assumption is satisfied for realistic CA concentrations [29]. Since the CT signal is measured in Hounsfield units (HU) which is directly related to the X-ray attenuation coefficients, the CA concentration $c(t)$ can be calculated directly from eq. 2.17.

Relative Signal Enhancement In DCE-MRI, k depends on tissue type and location and the measured signal has a non-linear proportionality to the concentration, in particular at peak concentrations. Additional calibration measurement, e.g., of the coil sensitivity profile or the flip angle, would be required to determine k. As an alternative, $c(t)$ can be approximated by normalizing the signal change to the baseline via *relative signal enhancement* (RSE), defined as:

$$c(t) = \frac{R_{10}}{r_1} \frac{S(t) - S_0}{S_0}. \tag{2.18}$$

Here, R_{10} is the precontrast relaxation rate and r_1 is the specific *relaxivity* of the CA. In comparison to eq. 2.17, the scaling factor for RSE is independent of coil sensitivity and flip angle, instead, it depends on the measurement of R_{10} in tissue and the supplying artery which is feasible with standard MRI techniques.

Absolute MR units The most accurate approach to derive CA concentrations uses the exact signal equation of the respective MR sequence to determine time-dependent relaxation rates $R_1(t)$, which have a linear relation to the CA concentration [30]. Due to the similarity to HU in DCE-CT regarding the linearity to concentration, the name *absolute MR units* has emerged [16]. As an example, the signal equation for the spoiled gradient echo (SPGR) sequence in steady state is used with the baseline signal $S_0(t)$, flip angle α and repetition time T_R:

$$S(t) = S_0 \sin(\alpha) \frac{1 - e^{-T_R R_1(t)}}{1 - \cos(\alpha) e^{-T_R R_1(t)}}. \tag{2.19}$$

Eq. (2.19) is solved for the time-dependent relaxation rate $R_1(t)$:

$$R_1(t) = -\frac{1}{T_R} \log\left(\frac{1 - A}{1 - \cos(\alpha) A}\right), \tag{2.20}$$

with the auxiliary variable

$$A = \frac{S(t)}{S_0} \frac{1 - e^{-T_R R_{10}}}{1 - \cos(\alpha) e^{-T_R R_{10}}}. \tag{2.21}$$

The time-dependent concentration $c(t)$ can then be calculated from the linear relation to the change in relaxation rates $R_1(t) - R_{10}$:

$$c(t) = (R_1(t) - R_{10})/r_1. \tag{2.22}$$

2.3.3 Tracer-kinetic Modeling

Measurement of DCE signal intensities and conversion to CA concentration builds the basis for the second step of DCE image analysis. The aim is to quantify contrast agent kinetics and derive hemodynamic information about the tissue of interest. The simplest approach is to describe the form of the concentration-time curves alone, e.g. maximum or time-to-peak, which yields *descriptive* parameters that lack clear physiological interpretation and are therefore only of limited use. Tracer-kinetic theory aims to describe the exact form of the concentration-time curves by considering the hemodynamic properties of the tissue of interest and the CA concentration in its supplying artery.

Tracer-kinetic Theory The standard tracer-kinetic theory is valid for tissues that are linear and stationary, i.e., the response is proportional to CA concentration and independent of the injection time [13, 14]. If satisfied, the CA concentration in the tissue $c_t(t)$ is related to the concentration in the arterial blood plasma $c_a(t)$ by the convolution with a *residual function* $R(t)$:

$$c_t(t) = F_p \cdot R(t) * c_a(t). \tag{2.23}$$

In eq. 2.23, $*$ denotes the convolution, F_p is the plasma flow that transports CA to the tissue and $c_a(t)$ represents the AIF. $R(t)$ is the monotonously decreasing tissue residue function that is always positive, satisfies $R(0) = 1$ and represents the fraction of tracer that is present in the tissue after an idealized, instantaneous (delta-shaped) CA administration.

The quantification of hemodynamic parameters thus relies on the determination of the product $F_p \cdot R(t)$, referred to as *impulse response function* (IRF), from the CA concentrations measured in tissue and the feeding artery. Since all hemodynamic information is contained in the IRF, the aim of tracer-kinetic modeling is to separate out the distinct contributions and derive appropriate parameters. A straight-forward approach is *deconvolution* [31], which has the purpose to numerically determine $R(t)$ from the measured concentrations without making assumptions about the form of the residue function. Since $R(0) = 1$, the maximum of the IRF can thus be determined as the plasma flow F_p and the area under $R(t)$ is the mean transit time (MTT).

Tissue Models Beyond the model-free deconvolution approach, model-based approaches exist that make explicit assumptions about the form of $R(t)$ and encode prior information about the tissue physiology. The internal structure and physiological process of microcirculation for typical tissues of interest is illustrated in Fig. 2.8.

Tissue is composed of cells, surrounded by the interstitium or extravascular extracellular space (EES) with volume v_e. The intravascular plasma volume of the capillaries is defined by the corresponding volume v_p. The plasma flow F_p pushes red blood cells and blood plasma through the capillary and transports CA into the system. It is defined as plasma volume (ml) that enters a unit of tissue volume (ml) per unit of time (min), resulting in units of 1/min. PS or permeability-surface area product is the rate of exchange between the blood plasma and the EES, defined as the number of contrast agent molecules (mmol) that extravasate per unit of time (min), plasma concentration (M) and tissue volume (ml). Like the plasma flow, PS has units 1/min. Furthermore, it is assumed that the backflux of PS equals the influx. Hence, the microcirculation in tissue is described by four independent parameters: F_p, PS, v_p and v_e.

Two Compartment Exchange Model Since most contrast agents in DCE imaging do not cross cell membranes and are therefore extracellular, most tracer-kinetic tissue models characterize two areas and their interactions with each other and the arterial system: the vascular space with the plasma volume v_p and the extravascular extracellular space with the interstitial volume v_e. The total measured tissue concentration $c_t(t)$ is thus composed as the weighted sum of the individual contributions:

$$c_t(t) = v_p c_p(t) + v_e c_e(t). \tag{2.24}$$

These two areas are assumed to be composed of interchanging, well-mixed *compartments*, in which the CA is distributed instantly and uniformly. Considering the conservation of mass in each compartment and that the outflux of a compartment is proportional to CA concentration, a set of differential equations can be built, which models the time-dependent concentrations in the compartments:

$$v_p \frac{\mathrm{d}c_p}{\mathrm{d}t}(t) = F_p c_a(t) - F_p c_p(t) + PS c_e(t) - PS c_p(t) \tag{2.25}$$

$$v_e \frac{\mathrm{d}c_e}{\mathrm{d}t}(t) = PS c_p(t) - PS c_e(t). \tag{2.26}$$

The vascular compartment is fed by two sources: the influx of arterial plasma concentration carried by the plasma flow $F_p c_a(t)$ and the influx from the concentration in the interstitium near the capillary wall carried by the permeability-surface area product $PS c_e(t)$. Concentration can drain off the compartment via extravasation to the interstitium by $PS c_p(t)$ or the outflux to the venous system by $F_p c_p(t)$. The in-

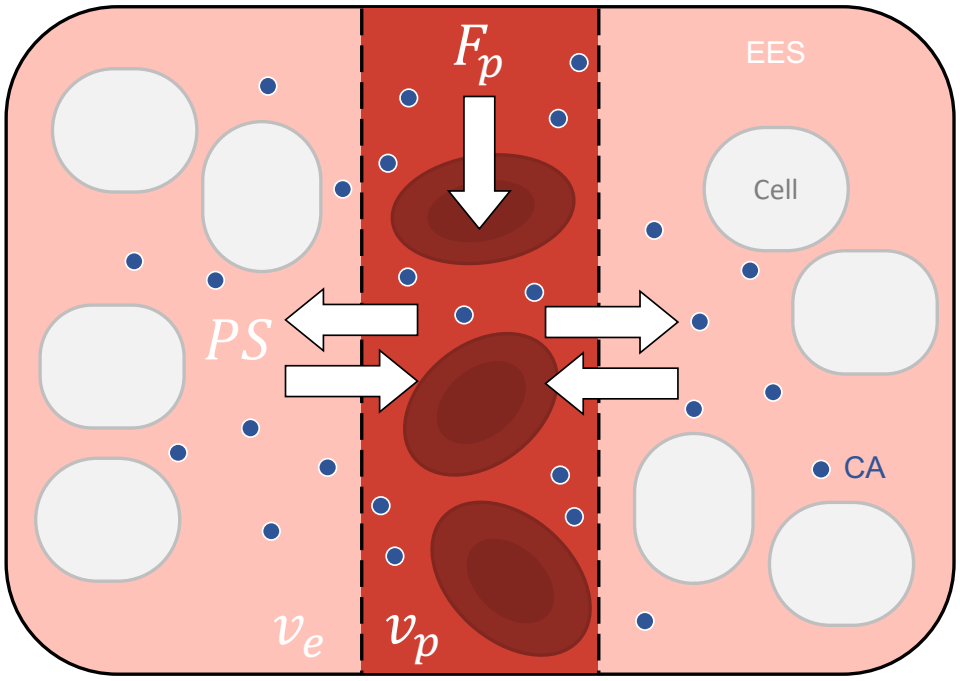

Figure 2.8: Schematic illustration of the internal structure and microcirculation in typical tissue of interest. The plasma flow F_p pushes red blood cells through a capillary and transports contrast agent (CA) into the system. The blood plasma with volume v_p exchanges solved molecules via the permeability-surface area product PS with the extravascular extracellular space (EES) with volume v_e. Adapted from [16].

terstitial compartment is assumed to only exchange concentration with the vascular compartment via the influx $PSc_p(t)$ and the outflux $PSc_e(t)$.

The so-called *two compartment exchange model* (2CX) is fully defined by eqs. 2.25 and 2.26 and the solution is a residue function $R_t(t) = R_p(t)+R_e(t)$, composed of the sum of two exponentials which yields the four model parameters F_p, PS, v_p and v_e. The 2CX model was developed for the analysis of the hemodynamics of tumors [32] but can be applied for most tissues where the contrast agent accesses the vascular and interstitial space.

Tofts Model A special case of the 2CX model is the widely used Tofts Model (TM) [33, 34], illustrated in Fig. 2.9. It assumes a negligible amount of intravascular tracer ($v_p \to 0$) and is valid only in tissue that is weakly vascularized [35]. By assuming negligible intravascular tracer volume, also the dispersion of CA within the vascular compartment can be ignored and the vascular plasma concentration is essentially replaced by the arterial plasma concentration. In order to fully define the TM, only a modified version of differential eq. 2.26 needs to be considered and the solution is

a monoexponential residue function $R_t(t)$. The total tissue concentration $c_t(t)$ can then be expressed, using eq. 2.23, as:

$$c_t(t) = K^{\text{trans}} e^{-tk_{\text{ep}}} * c_a(t). \tag{2.27}$$

Here, the transfer constant K^{trans} and the rate constant k_{ep} are defined as

$$K^{\text{trans}} = EF_p \quad k_{\text{ep}} = K^{\text{trans}}/v_e. \tag{2.28}$$

The introduction of K^{trans} underlines that the extraction fraction E of tracer into the capillary bed, related to F_p and PS ([36]), and the plasma flow F_p cannot be measured separately in tissues where the TM is applied.

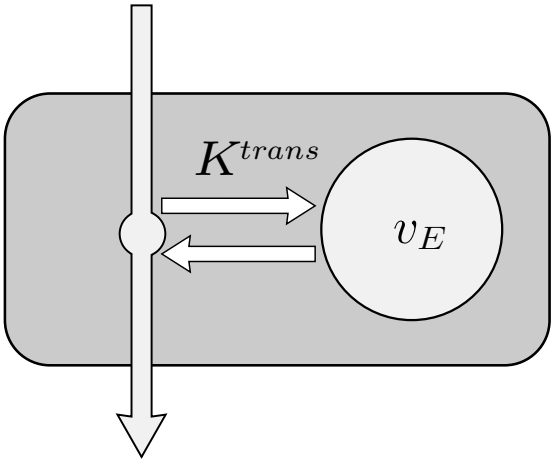

Figure 2.9: Schematic illustration of the Tofts model (TM) for tissue microcirculation which assumes a negligible amount of vascular volume v_p. The transfer constant K^{trans} describes the transportation and exchange of contrast agent (CA) with the extravascular extracellular space (EES), represented by the interstitial volume v_e.

2.4 DCE Analysis Methods

2.4.1 Conventional Tracer-kinetic Analysis

The standard method for DCE data analysis is to *fit* an appropriate tracer-kinetic model to the measured concentration-time curves to derive hemodynamic parameters. Fitting is usually performed by a non-linear least squares (NLLS) analysis ([37]) which aims to minimize the sum of squared residuals between model and data and yields, if successful, a *point estimate* of the model parameters. To account for noise in the measured data, an error term ϵ is added to the model equation $c_t(t)$.

The data y_i for one observation of tracer concentration at the timepoint t_i is then given as

$$y_i = c_t(t_i, \theta) + \epsilon_i. \tag{2.29}$$

where the model parameters are summarized in the vector θ.

The NLLS approach minimizes the sum of squared errors between the measured data y_i and the model function $c_t(t_i)$ for N timepoints, expressed as

$$\min \sum_{i=0}^{N} (y_i - c_t(t_i, \theta))^2 = \min \sum_{i=0}^{N} \epsilon_i^2. \tag{2.30}$$

Given successful convergence of the algorithm, the minimization yields the *best guess* parameters $\hat{\theta}$.

2.4.2 Bayesian Tracer-kinetic Analysis

Bayesian probabilistic modeling offers an alternative approach embedded in a framework of probability distributions. Briefly, a *prior* belief about the parameters of a tracer-kinetic model is formulated as a probability distribution. This way, domain expertise, e.g. physical constraints of model parameters, can be incorporated. Using dedicated algorithms, this prior belief is then updated with the information from measurements to infer the *posterior* probability distributions of the parameters given the data [38].

In detail, the *prior distribution* $P(\theta)$ reflects reasonable assumptions about the model parameters θ prior to the observation of data. A probability distribution is chosen for each parameter to meet physical or biological constraints such as upper or lower limits. Likewise to the conventional approach, the observational error ϵ_i for a measurement y_i at timepoint t_i is introduced (eq. 2.29) and furthermore assumed to be Gaussian with standard deviation σ. In the context of tracer-kinetic modeling, the *likelihood function* $P(y \mid \theta)$ can be constructed by computing the product of the probability distributions for each measurement given the true parameters [39]:

$$P(y \mid \theta) = \prod_{i=0}^{N} \mathcal{N}(y_i \mid c_t(t_i, \theta), \sigma^2), \tag{2.31}$$

with $\mathcal{N}$ representing a normal distribution and $c_t(t_i, \theta)$ the CA tissue concentrations.

The product of the likelihood function $P(y \mid \theta)$ and the prior distribution $P(\theta)$ is combined with the observed data to infer the joint posterior distribution $P(\theta \mid y)$ via Bayes' theorem:

$$P(\theta \mid y) = \frac{P(y \mid \theta)\, P(\theta)}{P(y)}. \tag{2.32}$$

The denominator in Eq. (2.32) is referred to as *model evidence* or *marginal likelihood* and calculates as $P(y) = \int P(\theta) P(y \mid \theta) \mathrm{d}\theta$ [40]. If the complexity of the model allows

no analytical solution to this integral, Markov Chain Monte Carlo (MCMC) methods ([41]) offer a means to determine the posterior probability distribution. Briefly, a MCMC algorithm draws samples from a target distribution, which equals the desired posterior distribution. The accepted parameter proposals are stored in a chain or trace of estimates ([42]).

An example of a Bayesian tracer-kinetic parameter estimation experiment is shown in Fig. 2.10. Left, a prior distribution $p(\theta)$ is illustrated for a single model parameter, representing domain knowledge such as $\theta \in [0, 1]$. The measured time-concentration curve y is displayed in the middle. Incorporating information about the measured data $p(y|\theta)$ and sampling via a MCMC algorithm infers the posterior distribution $p(\theta|y)$, displayed on the right. Hereby, information is gained about the parameter's accuracy and precision compared to the prior.

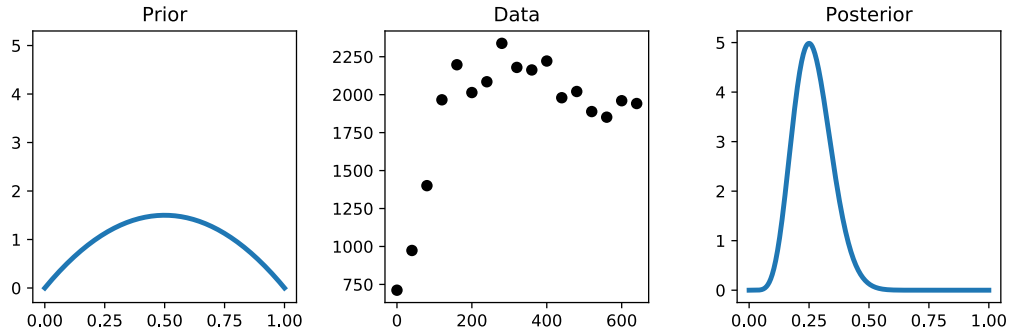

Figure 2.10: Bayesian modeling approach. Left: prior distribution $p(\theta)$ assumed for a parameter θ. Middle: observed data y. Right: Using the data distribution $p(y|\theta)$, the prior distribution $p(\theta)$ and sampling with an MCMC algorithm, the posterior distribution $p(\theta|y)$ is inferred. Information about the parameter's accuracy and precision is gained.

Through recent algorithmic developments [43] and the increasing availability of computational power, the use of Bayesian modeling approaches is spreading in various disciplines and has already shown to be a robust and accurate alternative for the analysis of MR imaging data [44–47]. The posterior probability distributions that result from Bayesian modeling greatly increase the interpretability of analysis results. Compared to simple NLLS point estimates, entire parameter probability distributions allow for a straightforward assessment of parameter changes. Hence, it is possible to assess if a parameter has truly changed in the course of a therapy or whether the change has only occurred within the uncertainty of the estimation ([48]).

2.4.3 Deep-Learning Approach for Perfusion Imaging

Deep learning (DL) is a class of machine learning methods with the capability to learn data-driven features from images. DL is particularly useful for analysis of large

bodies of unstructured data, such as radiological images and has been successfully applied to a variety of medical imaging task including image reconstruction, segmentation and image classification. Machine learning in general and DL in particular can be distinguished into three main paradigms which include supervised learning, unsupervised learning, and reinforcement learning. This work focuses on image classification, a supervised DL task. The aim is to learn a mapping function between input and output which is inferred from labeled training data, based on example input-output pairs. Based on [49, 50], this section briefly describes the basics of artificial neural networks and convolutional neural networks (CNNs) to understand their application to DCE imaging data. For a more in-depth introduction, the reader is advised to, e.g., [51].

Artificial Neural Networks An artificial neuron is the elementary unit in an artificial neural network and is schematically illustrated in Fig. 2.11. It takes multiple inputs x_i which have a separate weighting w_i plus an additional bias b. The artificial neuron computes the sum of weighted inputs

$$z = \sum_{i=1}^{m} w_i x_i + b \tag{2.33}$$

and passes it through a non-linear activation function $f(z)$ to generate an output $\hat{y}$.

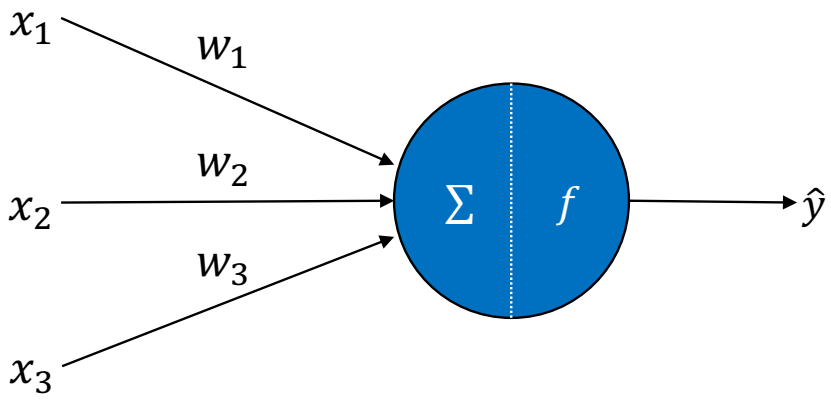

Figure 2.11: Schematic illustration of a simple artificial neuron. Given multiple weighted inputs $x_i w_i$, the artificial neuron calculates the sum, passes it through the activation function f and produces an output $\hat{y}$.

Activation functions are often nonlinear, monotonously increasing and differentiable. The activation function for the simplest form of an artificial neuron, the perceptron, is the Heaviside step function which produces a binary output. Sigmoid neurons are built similar to perceptrons but instead of a binary output, they yield any value

between zero and one. The output is calculated using the sigmoid function:

$$f(z) = \sigma(z) = \frac{1}{1 + e^{-z}} \tag{2.34}$$

Due to the characteristics of the sigmoid function, small changes in inputs x and weights w only cause small changes in the output $\hat{y}$. Furthermore, the derivative of the sigmoid function calculates as $\sigma(z)(1 - \sigma(z))$, a property that simplifies the learning process described in section 2.4.3. The sigmoid neuron is the basic building block of a simple artificial neural network.

The basic architecture of a simple neural network consists of multiple sigmoid neurons stacked in layers as illustrated in Fig. 2.12. The leftmost input layer consists of three input neurons, connected to the four neurons of the hidden layer in the middle. The rightmost output layer consists of a single output neuron. Since all neurons of a layer are connected to all neurons of its neighboring layers, this network is called fully connected (FC). A simple FC network with one layer is sufficient to represent any function [50].

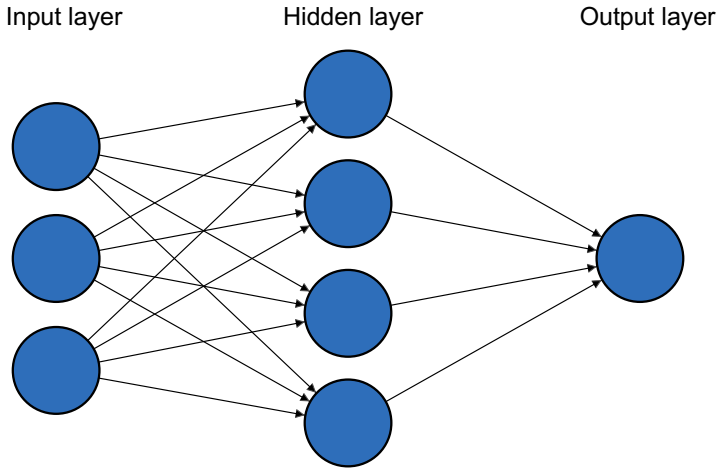

Figure 2.12: Architecture of a fully connected artificial neural network. Multiple sigmoid neurons are stacked in an input layer on the left, which is connected to a hidden layer in the middle and a single output layer on the right.

The Learning Process In order to learn, a neural networks requires a loss or cost function C, a quantitative measure which determines the differences between targets and predictions. The cost function takes weights and biases from all neurons in all layers into account. The aim of the learning process is to find weights and biases which minimize this cost function. For classification tasks, the binary cross-entropy cost function is often used which calculates the average cross-entropy across

all inputs:

$$C = -\frac{1}{n}\sum_{j=1}^{n}[y_j \ln a_j^L + (1-y_j)\ln(1-a_j^L)]. \tag{2.35}$$

Here, n is the number of inputs, y_j is the target output and $a_j^L = \sigma(z)$ the output or activation of all neurons in the output layer L. To find the minimum of C, small changes in the weights and biases of the network $\Delta\nu = (\Delta w, \Delta b)$ are introduced and the gradient $\nabla C = (\frac{\partial C}{\partial w}, \frac{\partial C}{\partial b})$ is calculated to relate these changes to the resulting changes in C:

$$\Delta C \approx \nabla C \Delta\nu. \tag{2.36}$$

The small changes in weights and biases can now be chosen to make ΔC negative, i.e., to reduce the cost function, by choosing an appropriate learning rate η:

$$\Delta\nu = -\eta\nabla C. \tag{2.37}$$

This process is known as gradient descent and is repeated until the cost function reaches a global minimum. To speed up training, stochastic gradient descent can be used, which approximates ∇C from a random sample of inputs.

Convolutional Neural Networks When dealing with images as network input, fully connected artificial neural networks as described above are not ideal, as the spatial structure of the image is lost when expressed in vector form. To take this spatial structure into account, convolutional neural networks are particularly well suited [49]. The pixel intensities of an image are then interpreted as spatially arranged input neurons. The basic building block of a CNN, illustrated in Fig. 2.13, relies on three key components: local receptive fields, shared weights and pooling.

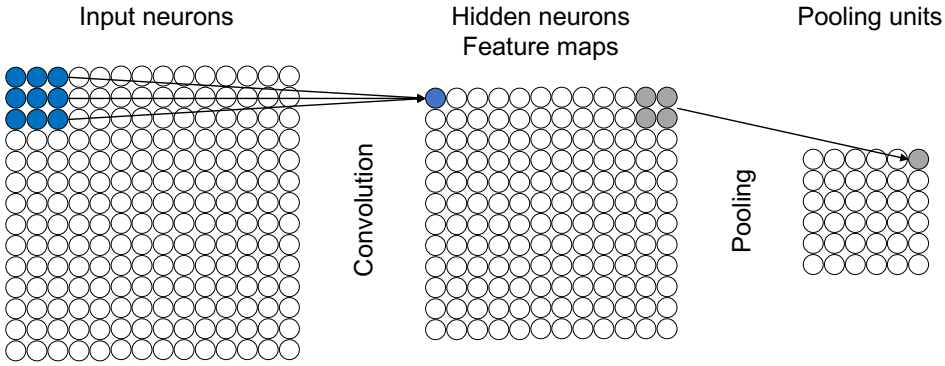

Figure 2.13: Basic building blocks of a convolutional neural network. Input neurons in the convolutional window with size 3×3 (blue) produce its corresponding hidden neuron in the middle. The kernel is moved across the image to produce the complete hidden layer. In a pooling step, the information of four neurons in the hidden layer (gray) is compressed to a single pooling unit on the right.

The local receptive field corresponds to a $l \times m$ window on the input neurons which connects the activation $a_{j,k}$ of those neurons to a hidden neuron. Each connection corresponds to a weight $w_{l,m}$ and the hidden neuron additionally learns a bias b; the calculation of the output of the j, k^{th} hidden neuron is defined as:

$$\sigma(b + \sum_{l} \sum_{m} w_{l,m} a_{j+l,k+m}). \tag{2.38}$$

Eq. 2.38 is known as convolution and σ is the sigmoid activation function as described above. By sliding this convolution window over the image, each local receptive field is connected to a neuron and the hidden layer is constructed. Each neuron in the hidden layer has the same weights and bias, i.e., the weights are shared across all neurons in one hidden layer, defining a feature map. A convolutional layer can have multiple feature maps responsible to detect different features, e.g., edges or horizontal lines, across the whole image.

A pooling layer usually follows the convolutional layer described above. Hereby, $n \times n$ neurons in the hidden layer are pooled, e.g., by taking the maximum or mean of the activations, and stored in the pooling layer. This way, information extracted by the feature maps can be further condensed to reduce the number of parameters for later layers. In other words, the exact location of a feature is less important for the network than its overall presence.

The architecture of a CNN typically consists of multiple of those building blocks of convolutional plus pooling layers. Finding the right architecture and its hyperparameters, e.g., size of the convolutional kernel, number of neurons to be pooled or total number of layers is ongoing research and depends strongly on the problem at hand. The convolutional layers were described for two dimensions above, however, they can be generalized to 3D when dealing with image volumes or simplified for 1D when dealing with sequential or temporal data.

The flexibility of CNN network design is especially useful for the analysis of time-resolved images, such as the perfusion data investigated in this work. In a supervised learning task, we aimed to predict a clinical endpoint directly from time-resolved DCE-CT images. For this purpose, a 2D CNN for spatial feature extraction was combined with a 1D CNN for temporal feature extraction. This unique architecture allows to extract information from time-resolved imaging data, otherwise only accessible through tracer-kinetic modeling. Other than tracer-kinetic modeling, this approach may permit to predict a clinical outcome directly from imaging data.

3 Contributions to the original publications

This chapter summarizes my contributions to the original articles upon which this work is based.

3.1 Contributions to publication I

The first publication (chapter 4) entitled *"Bayesian pharmacokinetic modeling of dynamic contrast-enhanced magnetic resonance imaging: validation and application"* was planned and designed in cooperation with all co-authors.

For the validation part of the study, I implemented and developed the Bayesian pharmacokinetic model in Python and the probabilistic programming framework *Stan* [52], in close cooperation with Michael Ingrisch. This included the implementation of a so-called prior predictive check to assess the credibility of the chosen prior distributions for the model parameters. As ground truth, I implemented and adapted simulated phantom data based on the quantitative imaging biomarkers alliance (QIBA) dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) phantom. I fitted the simulated data within a standard non-linear regression model and compared the results to the fits of the Bayesian model. To be able to compare the uncertainty estimates of the Bayesian model, I implemented a bootstrap approach which yields uncertainty estimates for the standard non-linear regression model. Then, I conducted the statistical analysis for the comparison of both approaches against the ground truth.

For the application part of the work, I cleaned and prepared a set of breast cancer DCE-MRI images from the cancer imaging archive. This included the conversion of signal intensities to contrast agent concentrations. I fitted the Bayesian model to imaging data of two visits and compared the derived imaging biomarker in order to assess treatment response. To this end, I created a script that uses Cohen's d to compare the probability distributions of the derived imaging biomarkers between the visits.

My results and interpretations for the validation and application part were critically discussed and reviewed with Michael Ingrisch and Olaf Dietrich. With the help of Michael Ingrisch, I drafted the manuscript which was read and approved by all co-authors.

3.2 Contributions to publication II

The second publication (chapter 5) entitled *"End-to-End Deep Learning Approach for Perfusion Data: A Proof-of-Concept Study to Classify Core Volume in Stroke CT"* was planned and designed in collaboration with all co-authors.

I implemented and developed a novel deep learning model in Python, using the dedicated framework *TensorFlow* [53], to predict dichotomized stroke core volume based on computed tomography (CT) perfusion data. Paul Reidler, Matthias Fabritius, Lars Kellert and Steffen Tiedt acquired and provided me with the CT perfusion data. First, two axial slices were selected based on the *Alberta stroke program early CT score* (ASPECTS) by Paul Reidler and Matthias Fabritius. Then, I preprocessed the resulting datasets which included resizing, interpolation, co-registration and standardization, by developing dedicated Python scripts. After intensive preparation, I trained the deep learning (DL) network and adapted hyperparameters and architecture details to optimize the predictions. I applied the final model to external test data which were preprocessed likewise to the in-house data to validate my results.

During the course of the development stage of the DL model, I had intensive discussions and re-evaluations of my findings and interpretations with Wolfgang G. Kunz, Michael Ingrisch, Paul Reidler and Matthias P. Fabritius who also helped me drafting the manuscript, as well as Birgit Ertl-Wagner, Balthasar Schachtner, Philipp Wesp and Olaf Dietrich. The final manuscript was read and approved by all co-authors.

4 | Publication I

IOP Publishing Phys. Med. Biol. 64 (2019) 18NT02 (11pp) https://doi.org/10.1088/1361-6560/ab3a5a

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
12 April 2019

REVISED
24 July 2019

ACCEPTED FOR PUBLICATION
12 August 2019

PUBLISHED
17 September 2019

NOTE

Bayesian pharmacokinetic modeling of dynamic contrast-enhanced magnetic resonance imaging: validation and application

Andreas Mittermeier[1,3], **Birgit Ertl-Wagner**[1,2], **Jens Ricke**[1], **Olaf Dietrich**[1] and **Michael Ingrisch**[1]

[1] Department of Radiology, Ludwig-Maximilians-University Hospital Munich, Munich, Germany
[2] Department of Medical Imaging, The Hospital for Sick Children, University of Toronto, Toronto, Canada
[3] Author to whom any correspondence should be addressed.

E-mail: Andreas.Mittermeier@med.lmu.de

Keywords: Bayesian inference, tracer-kinetic modeling, DCE-MRI, perfusion

Abstract

Tracer-kinetic analysis of dynamic contrast-enhanced magnetic resonance imaging data is commonly performed with the well-known Tofts model and nonlinear least squares (NLLS) regression. This approach yields point estimates of model parameters, uncertainty of these estimates can be assessed e.g. by an additional bootstrapping analysis. Here, we present a Bayesian probabilistic modeling approach for tracer-kinetic analysis with a Tofts model, which yields posterior probability distributions of perfusion parameters and therefore promises a robust and information-enriched alternative based on a framework of probability distributions. In this manuscript, we use the quantitative imaging biomarkers alliance (QIBA) Tofts phantom to evaluate the Bayesian tofts model (BTM) against a bootstrapped NLLS approach. Furthermore, we demonstrate how Bayesian posterior probability distributions can be employed to assess treatment response in a breast cancer DCE-MRI dataset using Cohen's *d*. Accuracy and precision of the BTM posterior distributions were validated and found to be in good agreement with the NLLS approaches, and assessment of therapy response with respect to uncertainty in parameter estimates was found to be excellent. In conclusion, the Bayesian modeling approach provides an elegant means to determine uncertainty via posterior distributions within a single step and provides honest information about changes in parameter estimates.

1. Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a noninvasive imaging technique used to quantify microvascular tissue perfusion with the help of a contrast agent (CA) (Ingrisch and Sourbron 2013). In MRI, a gadolinium-based CA is used most commonly and injected intravenously after the acquisition of pre-contrast baseline scans. The CA increases T1 and T2 relaxation rates of surrounding water protons and causes signal enhancement in a T1-weighted acquisition. By measuring multiple T_1-weighted images during the passage of the CA through the tissue of interest, a time-dependent CA concentration can be extracted from the signal-time course of each voxel. Besides determining semi-quantitative and descriptive parameters from the concentration curves, e.g. time to peak, area under curve, or maximum, quantitative perfusion parameters can be obtained by fitting pharmacokinetic (PK) models to the data (Roberts *et al* 2006, Sourbron and Buckley 2012, 2013). Popular PK models that characterize CA transport from DCE-MRI data are the classical Tofts model (TM) (Tofts 1997), the extended Tofts model and the two compartment exchange model (Sourbron and Buckley 2011).

The standard approach for estimating PK parameters from DCE-MRI data is using non-linear regression to determine a maximum likelihood estimator by non-linear least squares (NLLS) analysis (Seber and Wild 2003). For this purpose, an optimizing algorithm aims to minimize the sum of squared residuals between model and data and yields, if successful, a *point estimate* of model parameters. The NLLS approach is widely used, and a number of software packages provide non-linear regression implementation of a range of PK models (Huang *et al* 2014a, Beuzit *et al* 2016). Bayesian probabilistic modeling, on the other hand, offers an alternative modeling approach within a framework of probability distributions. Briefly, a *prior* belief about model parameters is formulated as a probability distribution; this allows to incorporate domain expertise, e.g. physical constraints. With

dedicated algorithms, this prior belief is then updated with the measured data and yields the *posterior* probability distributions of the parameters given the data (McElreath 2015). Through recent algorithmic developments (Hoffman and Gelman 2011) and the increasing availability of computational power, the use of Bayesian modeling approaches is spreading in various disciplines and has already shown to be a robust and accurate alternative for the analysis of MR imaging data (Schmid *et al* 2006, Orton *et al* 2007, Woolrich *et al* 2009, Dikaios *et al* 2017, Tietze *et al* 2018, Hansen *et al* 2019). The posterior probability distributions that result from Bayesian modeling greatly increase the interpretability of analysis results. Compared to simple point estimates, entire parameter probability distributions allow a straightforward assessment of, e.g. whether a parameter has truly changed in the course of a therapy, or whether the parameter change has only occurred within the uncertainty of the estimation (Shukla-Dave *et al* 2018).

In the present manuscript, we investigated Bayesian tracer-kinetic modeling in the context of DCE-MRI. To this end, we implemented a Bayesian TM (BTM) with the purpose to (i) evaluate accuracy against a NLLS approach using a digital reference object, (ii) validate uncertainty estimates against a bootstrapped NLLS approach to assess the precision and (iii) demonstrate how Bayesian posterior probability distributions can be used to assess treatment response in a breast cancer DCE-MRI dataset.

2. Materials and methods

2.1. Signal conversion and pharmacokinetic models

In a typical DCE-MRI experiment, time-resolved signal intensity curves $S(t)$ are extracted voxel-wise from multiple T_1-weighted images. To derive quantitative information, the measured signal intensities need to be converted to CA concentration curves. For this purpose, the signal equation for the spoiled gradient echo (SPGR) sequence in steady state can be used with the baseline signal $S_0(t)$, flip angle α, repetition time T_R and relaxation rate $R_1(t)$ as:

$$S(t) = S_0 \sin(\alpha) \frac{1 - e^{-T_R R_1(t)}}{1 - \cos(\alpha) e^{-T_R R_1(t)}}. \tag{1}$$

One can solve equation (1) for the time-dependent relaxation rate $R_1(t)$:

$$R_1(t) = -\frac{1}{T_R} \log\left(\frac{1 - A}{1 - \cos(\alpha) A}\right), \tag{2}$$

with the auxiliary variable

$$A = \frac{S(t)}{S_0} \frac{1 - e^{-T_R R_{10}}}{1 - cos(\alpha) e^{-T_R R_{10}}}. \tag{3}$$

A time-dependent concentration can then be calculated from the linear relation to the change in relaxation rates during and before administration of CA, $R_1(t)$ and R_{10}, respectively:

$$c(t) = (R_1(t) - R_{10})/r_1, \tag{4}$$

with the specific *relaxivity* of the gadolinium-based CA r_1 (Pintaske *et al* 2006).

A standard approach for the analysis of concentration-time curves in DCE-MRI data is the TM (Tofts and Kermode 1991, Tofts 1997, Sourbron and Buckley 2011) which assumes a negligible amount of intravascular tracer and describes CA transportation as:

$$c_t(t) = K^{\text{trans}} e^{-t k_{\text{ep}}} * c_p(t). \tag{5}$$

Here, $c_t(t)$ is the time-dependent concentrations of CA in the tissue of interest; $c_p(t)$ is the concentration in the blood plasma of the tissue-feeding artery, often referred to as arterial input function (AIF). $c_t(t)$ and $c_p(t)$ are connected with a convolution, expressed as '$*$'. The parameter v_e is the volume fraction of the interstitium, the extravascular extracellular space (EES). K^{trans} is defined as the transfer constant of CA between blood plasma and EES. The rate constant $k_{\text{ep}} = K^{\text{trans}}/v_e$ is the ratio of the transfer constant to the EES (Tofts *et al* 1999, Sourbron and Buckley 2011).

The tissue concentration $c_t(t)$ can be calculated from the measured signal $S(t)$ with equations (1)–(4) using the relaxation time T_{10} in tissue via $R_{10} = 1/T_{10}$. Plasma concentration $c_p(t)$ in the AIF can be calculated likewise using the relaxation time T_{10} of blood and the additional transformation from blood to plasma concentration via the hematocrit *hct*:

$$c_p(t) = c_b(t) \cdot (1 - hct). \tag{6}$$

Table 1. Parameters for the conversion from signal to concentration.

	T_{10} (Tissue)	T_{10} (Blood) (ms)	r_1 (Lmmol^{-1} ms^{-1})	α (°)	T_R (ms)	T_E	*hct*
QIBA DRO[a]	1000 ms	1440	0.0045	30	5	—	0.45
QIN breast[b]	1666 ms[c]	1440	0.0045	10	6.2	2.9 ms	0.45

[a] Quantitative imaging biomarker alliance digital reference object *QIBA_v6_Tofts*.
[b] Quantitative imaging network breast cancer dataset.
[c] Personal communication with the author of Huang *et al* (2014a).

The standard TM is used in the following within a classical NLLS likelihood framework and a Bayesian framework to quantify perfusion in simulated and measured DCE-MRI data. To account for noise in any observed data, an error term is added to the TM from equation (5) and an observation i is given as

$$y_i = c_t(t_i, \theta) + \epsilon_i \tag{7}$$

where the model parameters K^{trans} and v_e are summarized in the vector θ.

2.2. Data

2.2.1. Validation: QIBA DCE-MRI phantom

To evaluate accuracy of estimates and compare results of different fitting approaches, a simulated phantom with known PK parameters was investigated first. The quantitative imaging biomarkers alliance (QIBA)[3] provides several freely available test images for DCE-MRI analysis, known as digital reference objects (DRO). These have been used previously to validate various fitting algorithms and analysis toolkits (Ortuño *et al* 2013, Smith *et al* 2015, Debus *et al* 2019). The noise-free *QIBA_v6_Tofts* version was chosen here. The DRO contains simulated DCE-MRI data generated with the standard TM in equation (5) for a study duration of $t = 660$ s with a temporal resolution $\Delta t = 0.5$ s. Tissue concentration-time curves $c_t(t)$ have been created for all combinations of $K^{\text{trans}} \in \{0.01,0.02,0.05,0.1,0.2,0.35\}$ min^{-1} and $v_e \in \{0.01,0.05,0.1,0.2,0.5\}$, filling a 10×10 pixel patch for each combination. Table 1 lists the parameters stated in the QIBA description[4], following QIBA's DCE MRI quantification profile[5] to convert signal intensities to concentrations (compare equations (1)–(4)). For a more realistic setting, complex Gaussian noise with standard deviation $\sigma = 0.2$ relative to the pre-contrast baseline signal S_0 was added to the original noise-free test data. No noise was added to the AIF for simplicity and to be able to reliably relate our results to published work of Smith *et al* (2015) and Ortuño *et al* (2013). Figure 1 shows a snapshot of the DRO signal intensities at $t = 100s$, the AIF and an exemplary voxel with parameters $K^{\text{trans}} = 0.2$ min^{-1} and $v_e = 0.2$, respectively.

2.2.2. Application: breast cancer DCE-MRI data

The quantitative imaging network (QIN) aims at improving quantitative imaging and does so by sharing data which was acquired as part of various QIN studies, collected in the cancer imaging archive (TCIA)[6] (Clark *et al* 2013). A set of breast cancer DCE-MRI data (Huang *et al* 2014b) in DICOM format acquired from 10 patients was used to demonstrate the performance of the BTM on clinical data. The dataset contains DCE-MRI measurements acquired before (visit 1) and during (visit 2) preoperative neoadjuvant chemotherapy (NACT), respectively. For three patients, pathologic complete response (pCR) was reported, the remaining seven patients were classified as non-pCR. In addition, the dataset includes a region of interest (ROI) per patient, drawn by an experienced breast radiologist. A sample-averaged AIF is provided as blood concentration $c_b(t)$ and was converted to plasma concentration $c_p(t)$ using equation (6). Signal intensities within the ROI were converted to tissue concentrations using equations (1)–(4). Parameters for the conversion are specified in table 1, further details can be found in the original work by Huang *et al* (2014a).

2.3. Models and analysis

2.3.1. Non-linear least squares approach with bootstrapping

The standard evaluation of DCE-MRI data is performed in a likelihood framework by fitting a non-linear regression model to the concentration-time curve in every voxel. The NLLS approach minimizes the sum of squared errors between measured data y_i at timepoint t_i for $i = 0, 1, ..., N$ and the model function $c_t(t_i)$ in equation (7)

[3] https://sites.duke.edu/dblab/qibacontent/
[4] https://sites.duke.edu/dblab/files/2015/05/Dynamic_v6_beta1_description_Rev1.pdf
[5] http://qibawiki.rsna.org/images/1/12/DCE-MRI_Quantification_Profile_v1.0.pdf
[6] https://wiki.cancerimagingarchive.net/display/Public/Collections

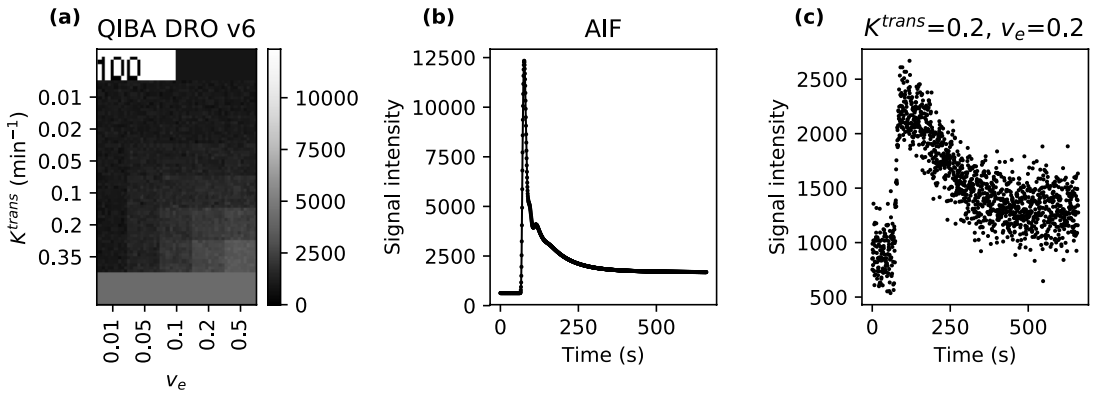

Figure 1. (a) Snapshot of the *QIBA_v6_Tofts* DRO at $t = 100$ s; the AIF is the bottom strip of the image, maximum of the AIF with timepoint labels in seconds are the top left strip and the zero patch ($K^{\text{trans}} = 0.0\,\text{min}^{-1}, v_e = 0.5$) is the top right strip. Intensity-time curves for AIF (b) and one pixel with $K^{\text{trans}} = 0.2\,\text{min}^{-1}$ and $v_e = 0.2$ with added noise (c), respectively.

$$\min \sum_{i=0}^{N} (y_i - c_t(t_i, \theta))^2 = \min \sum_{i=0}^{N} \epsilon_i^2 \tag{8}$$

to infer the *best guess* parameter $\hat{\theta}$. Assuming normally distributed noise ϵ_i, the least-squares estimator $\hat{\theta}$ equals the maximum-likelihood estimator (Seber and Wild 2003).

An implementation of the Broyden–Fletcher–Goldberg–Shano (L-BFGS) algorithm (Byrd *et al* 1994, Zhu *et al* 1997) in *SciPy*[7] (Jones *et al* 2001) was used for inference of the parameters via the *optimize.minimize* function. Initial values for K^{trans} and v_e were set to 0.001; constraints for K^{trans} and v_e were set to positivity and [0,1], respectively.The concentration curves of the DRO were then fitted and parameter maps were constructed for K^{trans} and v_e. By comparing them to the true parameter maps, percentage error maps were calculated as $\theta_{\%\text{err}} = (\hat{\theta} - \theta_{\text{true}})/\theta_{\text{true}}$.

A bootstrap method was implemented to assess the uncertainty of $\hat{\theta}$ (Kershaw and Buckley 2006). For that, the residuals, i.e. the difference between the fitted and the measured curve were calculated. In a next step, the residuals were resampled by randomly drawing samples with replacement. Subsequently, the resampled residuals were added to the fitted curve and the TM was used to determine another set of estimates, equivalent to inferring the original best guess. The number of iterations was set to 1000.

Uncertainty maps were then calculated from the bootstrap samples for the NLLS approach. Denoted as σ, half the width between 17th and 83rd percentile was considered a more robust measure for the precision than the standard deviation and is used throughout this work. For samples following a Gaussian normal distribution, σ would be equal to the standard deviation.

2.3.2. *Bayesian inference and implementation*

The alternative evaluation is performed in a Bayesian framework which infers a full posterior distribution $P(\theta \mid y)$ of the model parameters θ given an observation of data y. The observational error ϵ_i for each measurement y_i at timepoint t_i for $i = 0, 1, ..., N$ in equation (7) is assumed to be Gaussian with standard deviation σ. Hence, the joint observations of CA concentration in each voxel, conditional on the parameters, are modeled in the *likelihood* as

$$P(y \mid \theta) = \prod_{i=0}^{N} \mathcal{N}(y_i \mid c_t(t_i, \theta), \sigma^2), \tag{9}$$

with $\mathcal{N}$ representing a normal distribution and $c_t(t_i, \theta)$ the CA tissue concentration evaluated with the TM in equation (5).

Information about the parameters prior to the observation of data are specified in the *prior distribution* $P(\theta)$, enforcing physical or biological constraints. The likelihood of the data $P(y \mid \theta)$ and the product of the prior probability densities $P(\theta)$ are combined with the observed data to infer the joint posterior distribution via Bayes' theorem:

$$P(\theta \mid y) = \frac{P(y \mid \theta)\, P(\theta)}{P(y)}. \tag{10}$$

The denominator in equation (10) is referred to as model evidence and calculates as $P(y) = \int P(\theta) P(y \mid \theta) \mathrm{d}\theta$. If the complexity of the model allows no analytical solution to this integral, Markov Chain Monte Carlo (MCMC)

[7] Python 3.6.6, scipy 1.1.0, www.scipy.org/

methods (Gilks *et al* 1995) offer a means to determine the posterior probability distribution. Briefly, a MCMC algorithm draws samples from a target distribution, which equals the desired posterior distribution. The accepted parameter proposals are stored in a chain or trace of estimates (Kruschke 2014).

The BTM was implemented in *Stan* (Carpenter *et al* 2017), an open-source software package, using *pystan*[8]. In the present analysis, weakly informative priors were chosen which are applicable to a wide range of clinical DCE-MRI data without restrictions. Appendix provides a prior predictive check on these distributions, showing their weakly informative nature by comparing generated concentration curves with real observations. In particular, for the volume fraction $v_e \in [0, 1]$ a beta prior $v_e \sim \text{Beta}(\alpha = 2, \beta = 2)$ was chosen and for $K^{\text{trans}} \in \mathbb{R}_+$ a gamma prior was specified $K^{\text{trans}}(\text{min}^{-1}) \sim \text{Gamma}(\alpha = 1.1, \beta = 1/0.002)$. The prior for the standard deviation of the observational error was set to $\sigma(\text{mmol/L}) \sim \text{LogNormal}(\mu = 0, \sigma = 1)$. MCMC samples were drawn from the posterior distribution with the No-U-Turn (NUTS) algorithm (Hoffman and Gelman 2011). The number of iterations was set to 1000, sampled in two chains simultaneously, following a warm-up period of 500 iterations. Stan also reports divergences of the sampling algorithm and indicates the need to update the default settings of NUTS, e.g. initial step size and target acceptance rate.

To monitor the convergence of the MCMC chains to the target distribution, different diagnostics are automatically run alongside in Stan. The potential scale reduction statistic, $\hat{R}$, by Gelman and Rubin (1992) compares the sample variance within and across chains, and indicates if chains have not converged to a common distribution ($\hat{R} > 1.1$). The effective sample size N_{eff} indicates the degree of uncertainty in estimates due to autocorrelation of samples (Geyer 2011).

All concentration curves of the DRO were then fitted with the BTM to obtain posterior probability distributions of the parameters θ. To be able to compare the distributions to point estimates and to generate parameter maps, two hallmarks of the posterior distributions were determined: the median and, as for the bootstrap samples, half the distance between the 17th and 83rd percentile, denoted as σ. By comparing the median parameter maps to the true parameter maps, a map of the percentage error was calculated as above to assess the accuracy of estimates.

To evaluate the breast cancer DCE-MRI datasets, the mean tissue concentration curve over the ROI $c_{t,\text{ROI}}(t)$ was calculated for each patient and both visits. Subsequently, all concentration-time curves were fitted with the BTM to infer posterior distributions for the model parameters θ. To ensure that the model adequately captured the underlying data generating process, a *posterior predictive check* (PPC) was performed. Briefly, we used the BTM to generate new predictive data $\hat{y}$ and checked if it resembled the observed data. The full posterior distribution is exploited in this way to generate a *posterior predictive distribution*

$$P(\hat{y}|y) = \int P(\hat{y}|\theta)P(\theta|y)\text{d}\theta, \tag{11}$$

which propagates the uncertainty in the parameter estimates to uncertainty about prediction (Betancourt 2015, McElreath 2015, Gabry *et al* 2017). In this way, PPCs allow to detect systematic modeling errors and violations of model assumptions. Subsequently, the posterior distributions of K^{trans} were compared across visits for all patients with the objective to discriminate between patients with pCR and non-pCR.

2.4. Statistical analysis

A quantitative statistical measure for signal fidelity is the structural similarity index (SSIM) (Wang *et al* 2004). It gives an average value over similarities of three key elements of an image: luminance, contrast and structure (Wang and Bovik 2009). To assess the accuracy of parameter estimates for the DRO, the SSIM was calculated between the estimated and the true parameter maps. As a comparison, the root-mean-squared error (RMSE) was calculated alongside. In order to get reasonable values for RMSE, outliers in K^{trans}-estimates obtained from NLLS fitting needed to be restricted to one. In addition, the SSIM was calculated between the σ-uncertainty maps determined with the BTM and the bootstrapping method to assess similarities in the precision of estimates.

To compare the K^{trans} posterior distributions between visits for the breast cancer DCE-MRI dataset, Cohen's d was calculated for each of the ten patients as:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}}. \tag{12}$$

$\bar{x}$ represents the average K^{trans} value per visit, σ its standard deviation. In this way, the width of the posterior distributions are incorporated into a single value. Compared to just reporting the percentage change of K^{trans} mean values, the uncertainty in parameter estimation is accounted for. An univariate logistic regression (ULR) model, implemented in *scikit-learn*[9] (Pedregosa *et al* 2011), was fitted to the Cohen's d values. The receiver operating characteristic (ROC) area under curve (AUC) was calculated in order to obtain a quantitative measure for the assessment of response.

[8] Python 3.6.6, pystan 2.18.0, https://pystan.readthedocs.io/
[9] Python 3.6.6, scikit-learn 0.20.0, https://scikit-learn.org/

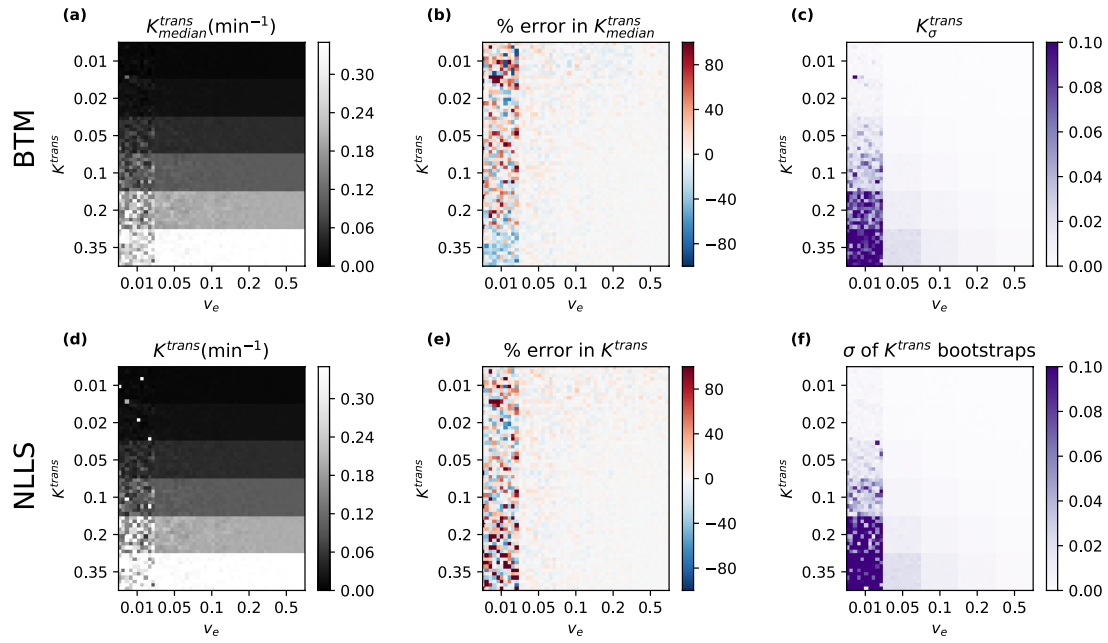

Figure 2. Estimated K^{trans} with BTM (top) and NLLS approach (bottom). The left column displays median (a) and point estimates (d). The middle column ((b) and (e)) shows the calculated percentage error between estimates and ground truth. The right column illustrates the uncertainty σ of the Bayesian posterior (c) and the additional bootstrap samples (f).

Table 2. SSIM and RMSE between estimated DRO K^{trans} and v_e parameter maps and ground truth for both approaches; SSIM of 100% indicates perfect similarity.

	K^{trans}		v_e	
	BTM (%)	NLLS (%)	BTM (%)	NLLS (%)
SSIM	96	91	92	94
RMSE	2.5	7.0	4.1	5.4

BTM = Bayesian tofts model; NLLS = Non-linear least squares approach; SSIM = Structural similarity index; RMSE = Root-mean-squared error.

3. Results

3.1. Validation: QIBA DCE-MRI phantom

Concentration-time curves of the DRO were evaluated within a Bayesian and likelihood framework. The resulting parameter estimates for K^{trans} are exemplarily shown in figure 2 for the Bayesian approach (a) and the NLLS reference (d). Note that the voxels in the Bayesian framework show median values of their respective posterior distributions while voxels in the likelihood framework represent point estimates. In general, the parameter maps show high accordance with the true values. The corresponding percentage error maps in the middle column (b) and (e) display relatively low errors for all regions with $v_e > 0.01$ for both methods. Low accuracy, hence high percentage errors are observed for regions where $v_e = 0.01$. SSIM between estimated and true K^{trans}-maps is higher for the BTM than for the NLLS approach. Furthermore, RMSE is lower for the BTM for both PK parameter maps. Details are provided in table 2.

The right column of figure 2 displays the precision of the parameter estimates evaluated with the BTM (c) and a bootstrapping method applied to the fit results of the NLLS approach (f). The visual analysis of the uncertainty maps reveals very similar patterns for both approaches, supported by a SSIM of 91%. The highest uncertainty occurs in regions with the highest percentage error for the fitting parameter estimates. The remaining parameter combinations have much greater precision. Information about divergences (BTM) and pixels where the NLLS algorithm did not find a solution can be found in table 3, together with computational times for fitting all 3000 pixels with BTM and NLLS approaches and the additional bootstrap analysis.

3.2. Application: breast cancer DCE-MRI data

Figure 3 shows representative signal intensity-time curves with the associated PPCs (a)–(c) and their corresponding K^{trans} posterior distributions (d). Here, the dark line illustrates the median and the increasingly lighter bands are the 20%, 40%, 60% and 80% highest density intervals (HDI) between the corresponding (0.4,0.6), (0.3,0.7), (0.2,0.8) and (0.1,0.9) percentiles of the posterior predictive distribution. The PPC in (a)

Table 3. Fitting process and parameter estimation of all 50×60 DRO concentration curves.

	BTM	NLLS
Divergences	17	27
Computational time: fitting	~48 min	~2 min
Computational time: uncertainty	*included*	~2100 min[a]

BTM = Bayesian tofts model; NLLS = Non-linear least squares approach.
[a] Based on additional bootstrap analysis.

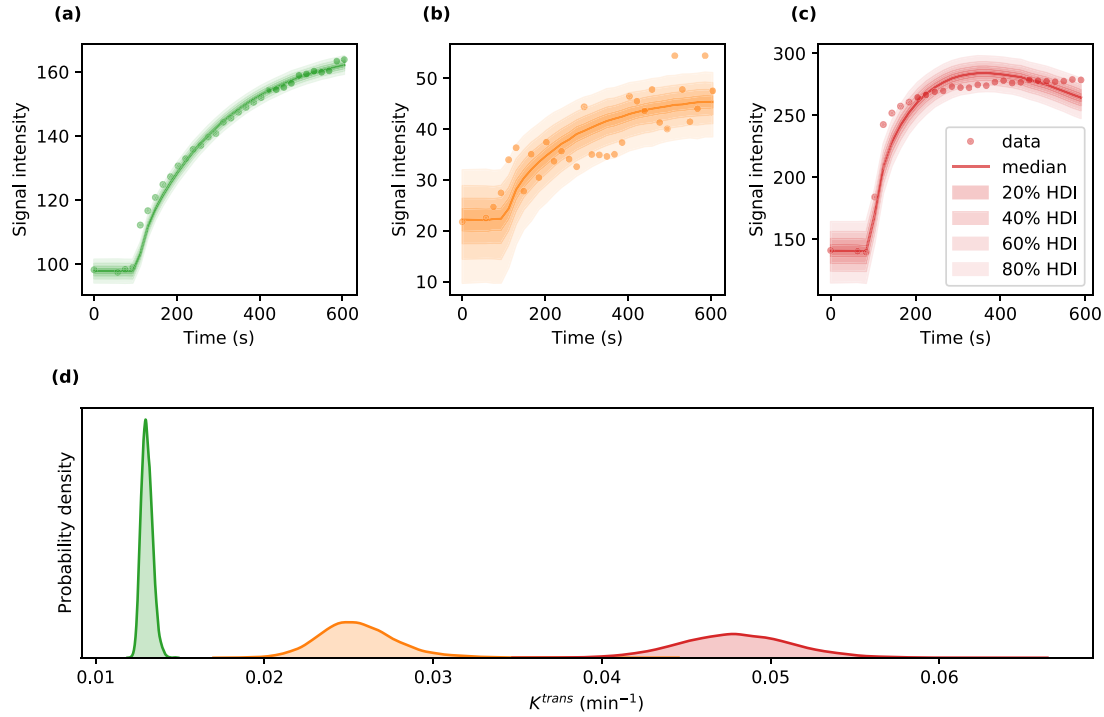

Figure 3. Observed signal intensity-time courses (dots) with posterior predictive distribution median (line) and the 20%, 40%, 60% and 80% highest density intervals (HDI); (a) for a good fit to data with low noise level, (b) for a good fit to data with high noise level, and (c) for a bad fit to data with low noise level. (d) Posterior distributions for K^{trans} estimated from the respective signal intensity-time curves (color-coded).

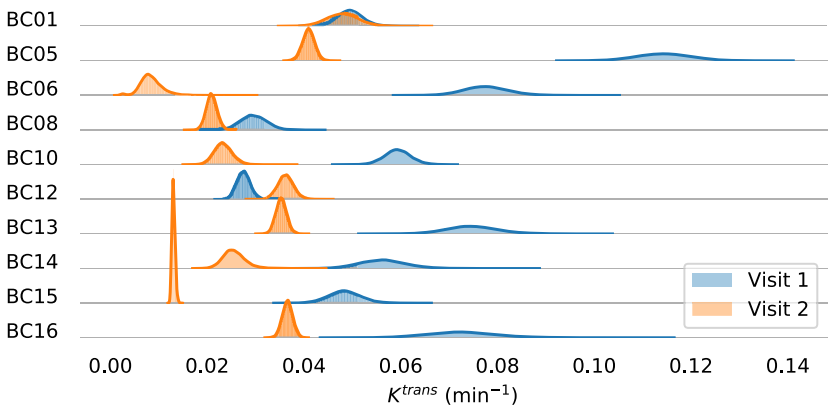

Figure 4. Posterior probability densities of K^{trans} for all patients. Blue corresponds to visit 1, orange to visit 2. BC05, BC06 and BC15 are labeled pCR, the rest non-pCR.

indicates a good fit of the model to the data, the corresponding posterior distribution (green) for K^{trans} is narrow. The PPC in (b) suggests that the chosen model provides a good fit to the data, the high noise level in the data is associated with a broader posterior distribution (orange). In (c), the noise level of the data is comparable to (a), however the PPC indicates a modeling error.

Figure 4 shows the posterior distributions of K^{trans} for all patients for visit 1 (blue) and visit 2 (orange), before and during NACT, respectively. With one exception, a general decrease in K^{trans} is observed. The degree

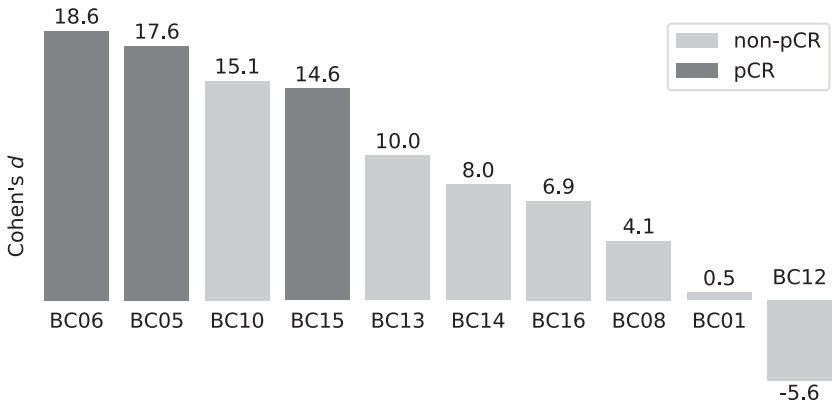

Figure 5. Cohen's *d* calculated from K^{trans} posterior distributions for all patients; sorted by value.

of change, dependent on the width of the posterior distributions, is summarized in Cohen's *d* values and visualized in figure 5; light-gray represents non-pCR, dark-gray pCR. The ULR analysis revealed a ROC AUC of 0.952. Computational time for fitting all 20 ROI-averaged concentration curves was ∼20 s for the BTM.

4. Discussion

In this study, we assessed posterior probability distributions of tracer-kinetic parameters obtained with a BTM against a standard NLLS approach. Validation with a DRO revealed high accuracy of BTM and NLLS approaches, indicated by strong similarity between estimated and ground truth maps. In addition, precision of estimates, assessed via the width of the posterior probability distributions and bootstrapping, respectively, was in very good agreement between both approaches. Analysis of the breast cancer DCE-MRI dataset with the BTM revealed that the degree of decrease in K^{trans} gives information about the pathologic response to NACT. The response in dependence of the uncertainty of parameter estimates was quantified with Cohen's *d*, calculated from the posterior distributions between visit 1 and 2. ULR modeling indicated excellent prediction of response.

Concerning the analysis of the DRO with the BTM, median parameter estimates were compared to the ground truth to assess the accuracy, otherwise not available with measured data. It was found that the Bayesian estimates generally have a very strong similarity with the ground truth, validating the accuracy of our BTM. The recovered parameters also have complementary regions of high and low percentage errors compared to the established NLLS fitting routine. RMS errors were lower for both implementations in the present work compared to similar DRO analysis by Smith *et al* (2015) and Ortuño *et al* (2013). Albeit, the results are in good comparison. Caution is still required for voxels with low v_e. Concentration curves with these parameter combinations have very limited intensity changes which practically vanish in the added background noise.

The variance of estimates inherent in the Bayesian posterior distribution was compared to a bootstrapping error analysis, performed likewise to the work of Kershaw and Buckley (2006). It was demonstrated that the uncertainty maps of the BTM resemble those calculated with the bootstrap analysis, validating the precision of parameters recovered with the BTM. To the best of our knowledge, only (Schmid *et al* 2006) implemented a Bayesian PK model with the objective to make use of the posterior probability distribution. Parts of the present study build up on their work and go beyond by capturing even more information from the posterior distribution. Their approach was applied to patient data only, whereas in the present work, accuracy and precision of estimates were validated with a digital phantom first.

Furthermore, we applied the BTM to the breast cancer DCE-MRI data, performed PPCs and investigated the posterior distributions. For a PPC, the observed data was compared to the posterior predictive distribution, illustrated as percentile intervals of highest density. A good fit to the data results in a posterior distribution which reflects the noise level in the data; low noise corresponds to a narrow posterior and vice versa. However, a bad fit to the data results in a broad posterior distribution despite a low noise level. This indicates a systematic modeling error which influences the information we gained about uncertainty. More complex PK models which incorporate additional assumptions about CA transport, e.g. the extended Tofts model, could be able to produce a better fit to certain data. Hence, assessing posterior distributions requires to check the corresponding data and fit before drawing any conclusions from it. While feasible for ROI-based analysis with only a handful of concentration curves, visual assessment is not possible in a pixelwise analysis. However, this can be necessary considering that tumorous tissue can be very heterogeneous (Wu *et al* 2018). An automated Bayesian model selection step as proposed in the work of Duan *et al* (2017) could be an effective means to reduce systematic modeling error but is beyond the scope of this study.

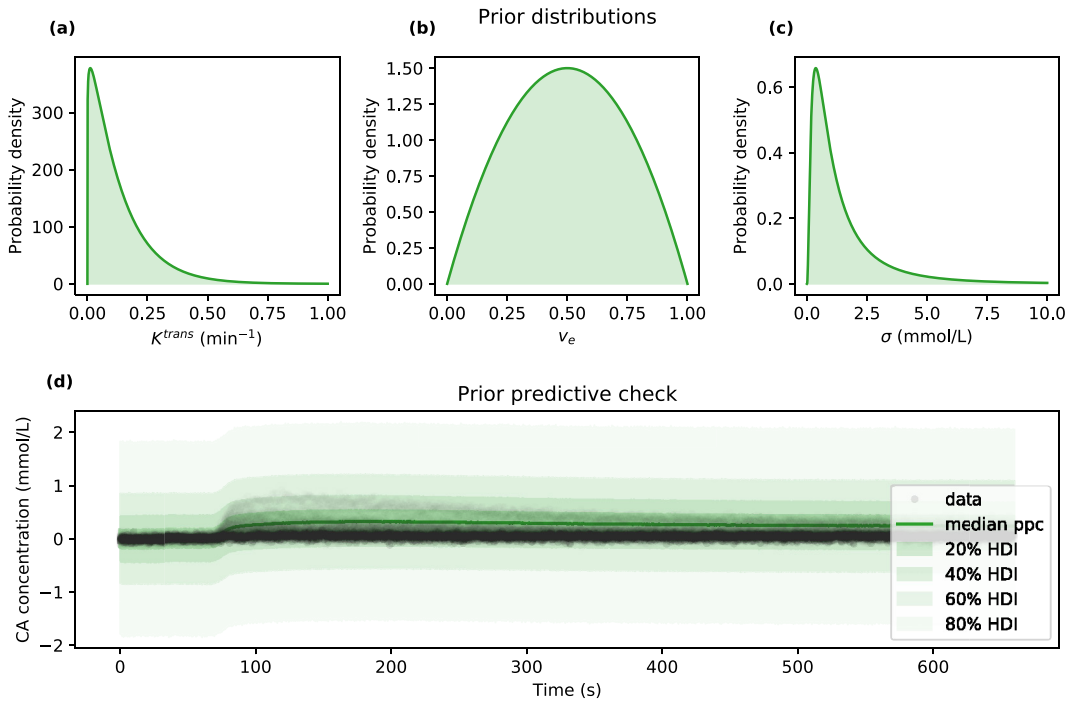

Figure A1. Prior probability density functions for K^{trans} (a), v_e (b), σ (c) and the subsequent prior predictive check (d). Simulated concentration-time curves (dots) for each parameter combination of the QIBA phantom are overlaid on the prior predictive distribution; illustrated by median (line) and the 20%–80% highest density intervals (HDI).

In order to assess therapy response for the patients in the breast cancer DCE-MRI dataset, Huang *et al* (2014a) showed in their original work that using visit 2 K^{trans} or the percentage change of K^{trans} between visits as metrics yields good to excellent results. However, the uncertainty in estimating PK parameters with tracer-kinetic models is not accounted for. For this purpose, we calculated Cohen's d as a means of quantitative change in parameter estimates which depends on the precision of estimates. In contrast, the authors of Schmid *et al* (2006) used the posterior distributions in each voxel to apply probabilistic thresholding, and then compared the mean values of K^{trans} between visits. Using Cohen's d metric, the assessment of response was found to be excellent by means of an ULR analysis. Considering the findings of the PPCs, including a model selection step as explained above could decrease the influence of systematic modeling errors on posterior distributions and hence Cohen's d values which may further improve assessment of therapy response.

Limitations of the present work include large computational time when fitting the BTM to the DRO-data. On the one hand, the MCMC sampling is time and memory consuming but necessary to avoid divergences. On the other hand, it yields a full posterior probability distribution with information about the uncertainty, and obtaining the same information with a bootstrap analysis of a NLLS fit requires even more computation time. Furthermore, the simulated DRO curves have a much higher time-resolution compared to measured data. Evaluating real DCE-MRI data increases the speed of the analysis greatly. Moreover, the influence of the chosen prior distributions on the results was not assessed in the present study.

In conclusion, we evaluated a BTM with a DRO, assessed accuracy and precision against the standard NLLS approach and showed how posterior distributions are used to assess therapy response. We demonstrated that Bayesian modeling provides an elegant means to assess posterior probability distributions, which are in good agreement with established approaches.

Acknowledgments

Funding: This work was supported by the research training group GRK 2274 of the DFG, Deutsche Forschungsgemeinschaft.

Appendix. Prior predictive check

To assess if the choice of prior distributions for the model parameters covers a reasonable range of concentration-time curves, it is useful to perform a *prior predictive check*. For this purpose, we generated 100 000 MCMC samples from the prior predictive distribution,

$$P(\hat{y}) = \int P(\hat{y} \mid \theta)\, P(\theta) \mathrm{d}\theta, \tag{A.1}$$

only considering the prior distributions without any actual data. This quantifies the range of possible observations $\hat{y}$, predicted by our model. In a prior predictive check, the predicted data is compared to real observations and the extent of extreme observations indicates the level of disagreement between domain expertise and model assumptions. Figure A1 shows the probability density functions of the chosen priors (a)–(c) and the prior predictive check (d). The black dots are actual observed data from the QIBA phantom, one curve for each parameter combination of K^{trans} and v_e, to assess the scope of possible phantom curves. The increasingly lighter green bands represent the 20%, 40%, 60% and 80% highest density intervals between the corresponding percentiles of the prior predictive distribution; the green line is the median thereof. We find that the model predicts observations that are more extreme than the phantom data but not too extreme to be unrealistic given the assumed observational error. Hence, we conclude that the chosen prior distributions are reasonable and weakly informative.

ORCID iDs

Andreas Mittermeier https://orcid.org/0000-0001-5743-473X
Olaf Dietrich https://orcid.org/0000-0001-6182-5039
Michael Ingrisch https://orcid.org/0000-0003-0268-9078

References

Betancourt M 2015 A unified treatment of predictive model comparison (arXiv:1506.02273 [stat])
Beuzit L, Eliat P A, Brun V, Ferré J C, Gandon Y, Bannier E and Saint-Jalmes H 2016 Dynamic contrast-enhanced MRI: study of inter-software accuracy and reproducibility using simulated and clinical data *J. Magn. Reson. Imaging* **43** 1288–300
Byrd R H, Lu P, Nocedal J and Zhu C 1994 A limited-memory algorithm for bound constrained optimization *SIAM J. Sci. Comput.* **16** 1190–208
Carpenter B, Gelman A, Hoffman M D, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P and Riddell A 2017 Stan: a probabilistic programming language *J. Stat. Softw.* **76** 1–32
Clark K *et al* 2013 The cancer imaging archive (TCIA): maintaining and operating a public information repository *J. Digit. Imaging* **26** 1045–57
Debus C, Floca R, Ingrisch M, Kompan I, Maier-Hein K, Abdollahi A and Nolden M 2019 MITK-ModelFit: a generic open-source framework for model fits and their exploration in medical imaging design, implementation and application on the example of DCE-MRI *BMC Bioinform.* **20** 31
Dikaios N, Atkinson D, Tudisca C, Purpura P, Forster M, Ahmed H, Beale T, Emberton M and Punwani S 2017 A comparison of Bayesian and non-linear regression methods for robust estimation of pharmacokinetics in DCE-MRI and how it affects cancer diagnosis *Comput. Med. Imaging Graph.* **56** 1–10
Duan C, Kallehauge J F, Bretthorst G L, Tanderup K, Ackerman J J H and Garbow J R 2017 Are complex DCE-MRI models supported by clinical data? *Magn. Reson. Med.* **77** 1329–39
Gabry J, Simpson D, Vehtari A, Betancourt M and Gelman A 2017 Visualization in Bayesian workflow *J. Royal Stat. Soc.* A **182** 389–402
Gelman A and Rubin D B 1992 Inference from iterative simulation using multiple sequences *Stat. Sci.* **7** 457–72
Geyer C J 2011 Introduction to Markov chain Monte Carlo *Handbook of Markov Chain Monte Carlo* ed S Brooks *et al* (London: Chapman and Hall) p 46
Gilks W R, Richardson S and Spiegelhalter D 1995 *Markov Chain Monte Carlo in Practice* (Boca Raton, FL: CRC Press)
Hansen M B, Tietze A, Haack S, Kallehauge J, Mikkelsen I K, Østergaard L and Mouridsen K 2019 Robust estimation of hemo-dynamic parameters in traditional DCE-MRI models *PLoS oNE* **14** e0209891
Hoffman M D and Gelman A 2011 The no-u-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo (arXiv:1111.4246)
Huang W *et al* 2014a Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge *Transl. Oncol.* **7** 153–66
Huang W, Tudorica A, Chui S, Kemmer K, Naik A, Troxell M, Oh K, Roy N, Afzal A and Holtorf M 2014b Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge *The Cancer Imaging Archive* (https://doi.org/10.7937/K9/TCIA.2014.A2N1IXOX)
Ingrisch M and Sourbron S 2013 Tracer-kinetic modeling of dynamic contrast-enhanced MRI and CT: a primer *J. Pharmacokinetics Pharmacodynamics* **40** 281–300
Jones E, Oliphant T and Peterson P 2001 SciPy: open source scientific tools for Python www.scipy.org/
Kershaw L E and Buckley D L 2006 Precision in measurements of perfusion and microvascular permeability with T1-weighted dynamic contrast-enhanced MRI *Magn. Reson. Med.* **56** 986–92
Kruschke J 2014 *Doing Bayesian Data Analysis* 2nd edn (New York: Academic)
McElreath R 2015 *Statistical Rethinking: a Bayesian Course with Examples in R and Stan* 1st edn (London: Chapman and Hall)
Orton M R, Collins D J, Walker-Samuel S, d'Arcy J A, Hawkes D J, Atkinson D and Leach M O 2007 Bayesian estimation of pharmacokinetic parameters for DCE-MRI with a robust treatment of enhancement onset time *Phys. Med. Biol.* **52** 2393
Ortuño J E, Ledesma-Carbayo M J, Simões R V, Candiota A P, Arús C and Santos A 2013 DCE@urLAB: a dynamic contrast-enhanced MRI pharmacokinetic analysis tool for preclinical data *BMC Bioinform.* **14** 316
Pedregosa F *et al* 2011 Scikit-learn: machine learning in Python *J. Mach. Learn. Res.* **12** 2825–30

Pintaske J, Martirosian P, Graf H, Erb G, Lodemann K P, Claussen C D and Schick F 2006 Relaxivity of gadopentetate dimeglumine (magnevist), gadobutrol (gadovist), and gadobenate dimeglumine (MultiHance) in human blood plasma at 0.2, 1.5, and 3 T *Investigative Radiol.* **41** 213–21

Roberts C, Issa B, Stone A, Jackson A, Waterton J C and Parker G J M 2006 Comparative study into the robustness of compartmental modeling and model-free analysis in DCE-MRI studies *J. Magn. Reson. Imaging* **23** 554–63

Schmid V J, Whitcher B, Padhani A R, Taylor N J and Yang G Z 2006 Bayesian methods for pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging *IEEE Trans. Med. Imaging* **25** 1627–36

Seber G A F and Wild C J 2003 *Nonlinear Regression* (New York: Wiley)

Shukla-Dave A *et al* 2018 Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials *J. Magn. Reson. Imaging* **49** e101–e121

Smith D S, Li X, Arlinghaus L R, Yankeelov T E and Welch E B 2015 DCEMRI.jl: a fast, validated, open source toolkit for dynamic contrast enhanced MRI analysis *PeerJ* **3** e909

Sourbron S P and Buckley D L 2011 On the scope and interpretation of the Tofts models for DCE-MRI *Magn. Reson. Med.* **66** 735–45

Sourbron S P and Buckley D L 2012 Tracer kinetic modelling in MRI: estimating perfusion and capillary permeability *Phys. Med. Biol.* **57** R1–33

Sourbron S P and Buckley D L 2013 Classic models for dynamic contrast-enhanced MRI *NMR Biomed.* **26** 1004–27

Tietze A, Nielsen A, Mikkelsen I K, Hansen M B, Obel A, Østergaard L and Mouridsen K 2018 Bayesian modeling of dynamic contrast enhanced MRI data in cerebral glioma patients improves the diagnostic quality of hemodynamic parameter maps *PLoS One* **13** e0202906

Tofts P S 1997 Modeling tracer kinetics in dynamic Gd-DTPA MR imaging *J. Magn. Reson. Imaging* **7** 91–101

Tofts P S and Kermode A G 1991 Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts *Magn. Reson. Med.* **17** 357–67

Tofts P S *et al* 1999 Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols *J. Magn. Reson. Imaging* **10** 223–32

Wang Z and Bovik A 2009 Mean squared error: Love it or leave it? A new look at signal fidelity measures *IEEE Signal Process. Mag.* **26** 98–117

Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Proc.: Publ. IEEE Signal Proc. Soc.* 13 600–12

Woolrich M W, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M and Smith S M 2009 Bayesian analysis of neuroimaging data in FSL *NeuroImage* **45** S173–86

Wu J, Cao G, Sun X, Lee J, Rubin D L, Napel S, Kurian A W, Daniel B L and Li R 2018 Intratumoral spatial heterogeneity at perfusion MR imaging predicts recurrence-free survival in locally advanced breast cancer treated with neoadjuvant chemotherapy *Radiology* **288** 26–35

Zhu C, Byrd R H, Lu P and Nocedal J 1997 Algorithm 778: L-BFGS-B: fortran subroutines for large-scale bound-constrained optimization *ACM Trans. Math. Softw.* **23** 550–60

5 | Publication II

diagnostics

MDPI

Article

End-to-End Deep Learning Approach for Perfusion Data: A Proof-of-Concept Study to Classify Core Volume in Stroke CT

Andreas Mittermeier [1,*], Paul Reidler [1], Matthias P. Fabritius [1], Balthasar Schachtner [1,2], Philipp Wesp [1], Birgit Ertl-Wagner [3], Olaf Dietrich [1], Jens Ricke [1], Lars Kellert [4], Steffen Tiedt [5], Wolfgang G. Kunz [1] and Michael Ingrisch [1]

1 Department of Radiology, University Hospital, LMU Munich, 81377 Munich, Germany; paul.reidler@med.uni-muenchen.de (P.R.); matthias.fabritius@med.uni-muenchen.de (M.P.F.); balthasar.schachtner@med.uni-muenchen.de (B.S.); philipp.wesp@med.uni-muenchen.de (P.W.); olaf.dietrich@med.uni-muenchen.de (O.D.); jens.ricke@med.uni-muenchen.de (J.R.); wolfgang.kunz@med.uni-muenchen.de (W.G.K.); michael.ingrisch@med.uni-muenchen.de (M.I.)
2 Comprehensive Pneumology Center (CPC-M), German Center for Lung Research (DZL), 81377 Munich, Germany
3 Department of Diagnostic Imaging, The Hospital for Sick Children, University of Toronto, Toronto, ON M5G 1X8, Canada; birgitbetina.ertl-wagner@sickkids.ca
4 Department of Neurology, University Hospital, LMU Munich, 81377 Munich, Germany; lars.kellert@med.uni-muenchen.de
5 Institute for Stroke and Dementia Research, University Hospital, LMU Munich, 81377 Munich, Germany; steffen.tiedt@med.uni-muenchen.de
* Correspondence: Andreas.Mittermeier@med.uni-muenchen.de; Tel.: +49-89440074621

Abstract: (1) Background: CT perfusion (CTP) is used to quantify cerebral hypoperfusion in acute ischemic stroke. Conventional attenuation curve analysis is not standardized and might require input from expert users, hampering clinical application. This study aims to bypass conventional tracer-kinetic analysis with an end-to-end deep learning model to directly categorize patients by stroke core volume from raw, slice-reduced CTP data. (2) Methods: In this retrospective analysis, we included patients with acute ischemic stroke due to proximal occlusion of the anterior circulation who underwent CTP imaging. A novel convolutional neural network was implemented to extract spatial and temporal features from time-resolved imaging data. In a classification task, the network categorized patients into small or large core. In ten-fold cross-validation, the network was repeatedly trained, evaluated, and tested, using the area under the receiver operating characteristic curve (ROC-AUC). A final model was created in an ensemble approach and independently validated on an external dataset. (3) Results: 217 patients were included in the training cohort and 23 patients in the independent test cohort. Median core volume was 32.4 mL and was used as threshold value for the binary classification task. Model performance yielded a mean (SD) ROC-AUC of 0.72 (0.10) for the test folds. External independent validation resulted in an ensembled mean ROC-AUC of 0.61. (4) Conclusions: In this proof-of-concept study, the proposed end-to-end deep learning approach bypasses conventional perfusion analysis and allows to predict dichotomized infarction core volume solely from slice-reduced CTP images without underlying tracer kinetic assumptions. Further studies can easily extend to additional clinically relevant endpoints.

Keywords: CT perfusion; stroke; deep learning; contrast-enhanced perfusion imaging; convolutional neural networks; end-to-end modeling

check for updates

Citation: Mittermeier, A.; Reidler, P.; Fabritius, M.P.; Schachtner, B.; Wesp, P.; Ertl-Wagner, B.; Dietrich, O.; Ricke, J.; Kellert, L.; Tiedt, S.; et al. End-to-End Deep Learning Approach for Perfusion Data: A Proof-of-Concept Study to Classify Core Volume in Stroke CT. *Diagnostics* **2022**, *12*, 1142. https://doi.org/10.3390/diagnostics12051142

Academic Editor: David S. Liebeskind

Received: 8 March 2022
Accepted: 3 May 2022
Published: 5 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Acute ischemic stroke occurs when a blood clot interrupts the blood flow (perfusion) to the brain, most commonly in a supplying artery—this causes cell death in the hypoperfused areas [1]. Historically, cerebral perfusion imaging was performed using positron emission tomography using radioactive labeled oxygen to determine oxygen fraction and cerebral

metabolic rate for oxygen or single photon emission computed tomography. However, logistics and application of radiotracers made both modalities unfeasible for the emergency setting. Today, computed tomography perfusion (CTP) is the most frequently used method to classify the salvageable brain tissue (penumbra) from the irreversibly damaged core in order to support clinical decision-making [2–4].

CTP is based on consecutive sampling of cerebral tissue attenuation after intravenous bolus injection of an iodinated contrast agent. A time-attenuation curve in every voxel represents the passage of the contrast agent through the brain in the reconstructed 4D image. After conversion to concentration, tracer-kinetic analysis aims to quantitatively evaluate the time-concentration curves by estimating perfusion parameters, e.g., cerebral blood volume, cerebral blood flow, time to peak, and mean transit time [5]. The most common approach uses deconvolution: An arterial input function (AIF) is determined in a large feeding artery and the time-concentration curves are deconvolved voxel-wise with the AIF to estimate perfusion parameters [6–8].

Radiologists as human experts then examine the generated perfusion parameter maps to detect hypoperfused areas, i.e., penumbra and core, and decide among treatment options. In the setting of acute ischemic stroke, CTP can help to identify patients who have a large penumbra and a small core, as they are likely to have a favorable response to reperfusion therapies [9,10]. Additionally, it was shown that CTP can help identify stroke mimics like epilepsy [11] and improves the detection performance for peripheral ischemia with often minor clinical symptoms, which is paramount for future therapy concepts on medium vessel occlusion [12]. However, availability and usage of advanced stroke imaging methods, including CTP, vary considerably among sites and geographical areas, with only around half of centers using those methods frequently [13].

The value of convolutional neural networks (CNN) has been demonstrated for a variety of medical imaging tasks, e.g., image reconstruction, object detection, segmentation, or classification [14–19]. The reason for the success of these systems is based on the capability of CNNs to learn data-driven features from pixels directly. Multiple nonlinear processing layers produce a high-level representation of features in images. Consequently, CNN-based approaches were proposed and applied to perfusion imaging analysis. For dynamic contrast-enhanced (DCE) MRI for example, CNN models were developed to estimate perfusion parameters maps directly from the data without the requirement for a standard deconvolution process [20,21]. A recent study [22] proposed a voxelwise prediction of infarct status from stroke CTP, but with the use of additional clinical and tracer-kinetic related data.

Commercial CTP analysis approaches often require domain expertise—e.g., by localizing, verifying, or correcting vessels for the measurement of arterial input functions. Uncertainties, e.g., induced by partial volume effects, propagate into the calculation of perfusion parameter maps via tracer kinetic modeling and ultimately in the process of clinical decision-making. Deep learning, on the other hand, may enable the direct prediction of clinical endpoints from complex imaging data with minimal user input. Starting from the baseline model developed in this study, more complex models can be adapted to relevant clinical endpoints in acute ischemic stroke, e.g., grade of disability or quality of life. For ischemic stroke, the most widely applied measure for neurological outcome uses the modified Rankin scale at day 90 after stroke. Clinical endpoints are likely associated with subtle patterns in the data and development of deep learning models generally requires large datasets [23]. In a proof-of-concept study, we, therefore, aimed to categorize patients into small or large core with an end-to-end deep learning approach for slice-reduced CT perfusion data.

2. Materials and Methods

2.1. Study Population, Image Acquisition and Core Volumetry

In this retrospective study, we included a training cohort (n = 217) (Figure 1) from among 234 consecutive acute ischemic stroke patients with available raw CTP data from

a prospectively acquired cohort (German Stroke Registry, NCT03356392). All patients were treated with endovascular mechanical thrombectomy at our institution. We excluded patients with inconsistent CTP images that did not comply with a standardized time resolution of 1.5 s or a scan duration of 48 s. Patients underwent CTP on admission using a SOMATOM Definition Force, AS+, or Flash CT scanner (Siemens Healthineers, Forchheim, Germany). Automated calculation of ischemic core was performed using the CT vendor's proprietary software (syngo Neuro Perfusion CT; Siemens Healthineers, Forchheim, Germany), which applies a threshold cerebral blood volume of <1.2 mL/100 mL. Median core volume was calculated and used as threshold value for binary classification: Small core < median core, large core > median core. As a proof-of-principle, we used median core volume to ensure a balanced split of the training data, in contrast to a fixed threshold value of 70 mL, which can sometimes be found in the literature [24].

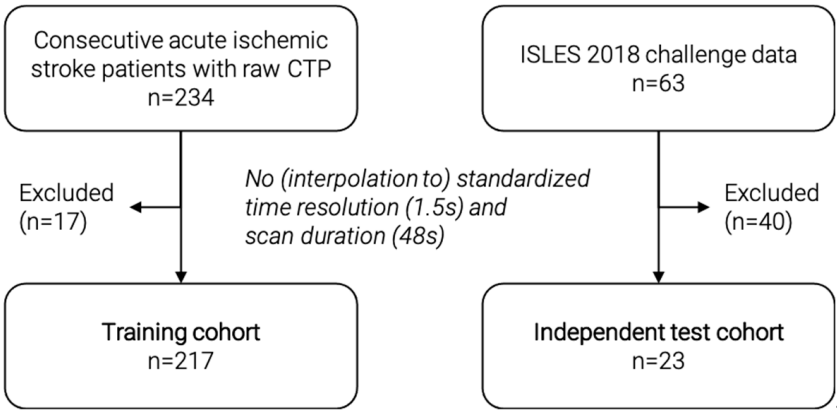

Figure 1. Flowchart of patient selection for the training and independent test cohort. CTP = CT perfusion.

For independent validation, we included a second external test cohort (n = 23) from among 63 patients of the external, publicly available ISLES 2018 challenge data set [25,26]. We excluded patients with inconsistent CTP images that could not be interpolated to the standardized time resolution of 1.5 s and a scan duration of 48 s. The challenge data include core segmentations, which were used to calculate the ischemic core volume. For this purpose, the number of voxels in the segmentation was multiplied by the voxel dimensions to get an estimated core volume in ml. The median core volume of the training dataset was used as threshold value for binary classification. Figure 1 shows a detailed flowchart of patient selection for both cohorts.

2.2. Preprocessing, Batch Generation, and Data Augmentation

Internal and external datasets were both preprocessed in two steps. First, two axial slices covering the middle cerebral artery territory (basal ganglia and supraganglionic level) based on the Alberta stroke program early CT score (ASPECTS) [27] regions were selected by radiologists. In a second fully automated step, the selected slices were resized and interpolated to 128 × 128 pixels in-plane resolution with 200 × 200 mm^2 length. All slices along the time axis were co-registered to the first slice at t = 0 to reduce motion artifacts. All data were processed using custom Python (version 3.8.5) [28] scripts including the publicly available packages SimpleITK (version 2.0.2) [29] and Scikit-learn (version 0.23.2) [30].

During training, a custom batch generator returned a random subset of samples (batch size = 12) from the complete dataset and normalized each batch to zero mean and unit variance. Online data augmentation was applied to each batch in the form of random rotation in the range of (−15°, 15°), xy-shift (−10 pixel, 10 pixel), and vertical flip (True, False) before passing it on to the network.

2.3. Network Architecture

The proposed network architecture was implemented in Python and TensorFlow (version 2.3.0) [31] and is illustrated in Figure 2. It consists of two submodels with identical

architecture for each of the selected axial CTP slices. The standardized and augmented 2D+t input images were fed into each submodel and processed through the pipeline to extract spatial and temporal features. The resulting features are concatenated, passed through a fully connected dense layer, and classified (Figure 2 top).

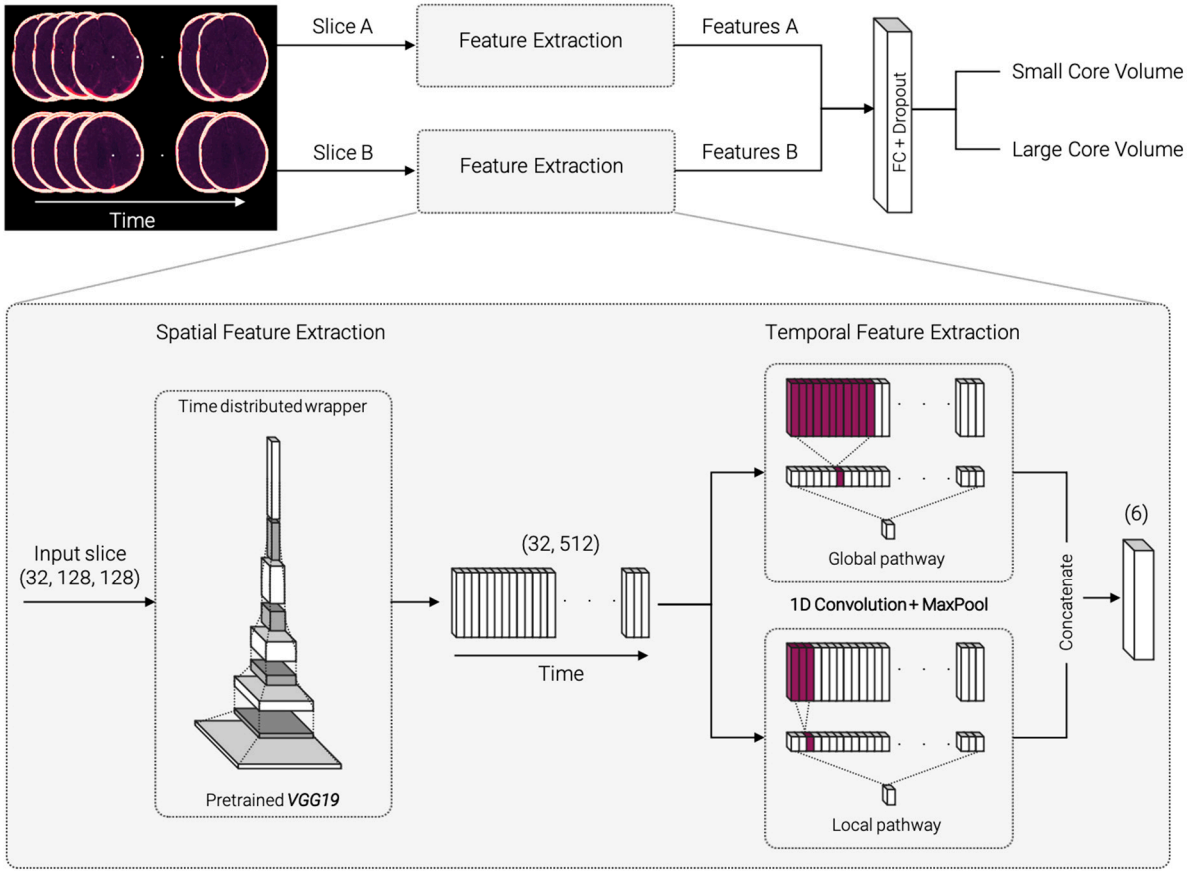

Figure 2. Model architecture overview and detailed, zoomed-in view of the spatial and temporal feature extraction process. The selected slices A and B are fed into identical submodels for spatial and temporal feature extraction. Spatial feature extraction consists of identical, pretrained VGG19 networks for each timepoint of the input images. The resulting feature vector is passed on to the temporal feature extraction. 1D convolutions with two different kernel sizes are carried out in a global and local pathway. The extracted features A and B for both submodels are concatenated, fully connected (FC), and classified.

Figure 2 (bottom, zoomed-in) displays a detailed demonstration of the feature extraction part. For spatial feature extraction, each 2D image on the time axis is fed into a VGG19 model [32], pretrained on 2D image net data [33]. The weights are shared across all timepoints within a "TimeDistributed" framework. The resulting 32×512 feature matrix is passed on to the temporal feature extraction step. The temporal feature extraction consists of a 1D convolution with three filters followed by a max pooling layer and is divided into a global and a local pathway. In the global pathway, the 1D convolution is performed with kernel size 11, in the local pathway with kernel size 3. This ensures that the model can capture both smaller and larger changes along the time course. The resulting feature vectors are concatenated and passed on to a dense layer with 32 units. Classification is performed using a sigmoid layer. The source code is made publicly available on the development platform Github (https://github.com/AndreasMittermeier/stroke-perfusion-CNN (accessed on 7 March 2022)).

2.4. Training, Validation and Testing

The proposed network was trained, validated, and tested using all included patients from the training cohort within a 10-fold cross-validation (CV). The dataset was randomly split into ten folds according to an 8:1:1 ratio of training, validation, and test. Eight folds were used for training the network. The number of training epochs was set to 500. The validation fold was used to evaluate the model after each epoch and stop training once the validation loss stopped decreasing for 200 epochs (patience = 200). After the last epoch, the test fold was evaluated by the model with the best weights, i.e., the weights which yielded the lowest validation loss. After ten CV iterations, each fold was used for unbiased testing once. The area under the receiver operating characteristic curve (ROC-AUC) was used as evaluation metric. Mean and standard deviation (SD) of ROC-AUC values were reported for the 10-fold CV.

In addition, we performed an ablation study on the effect of the local and global temporal feature extraction. To this end, we trained two reduced models using the (i) local feature extractor alone and using the (ii) global feature extractor alone. Training and evaluation on the test folds was performed in the same CV approach as described above. Mean and SD of the ROC-AUC values were reported and compared with those of the full model.

To evaluate the independent test cohort, an ensemble method was used. The final model was constructed by averaging the predictions from the ten models trained in the CV. The final model was applied to the independent test cohort and the ROC-AUC was reported.

3. Results

Two hundred seventeen patients were included in the training cohort and 23 patients in the independent test cohort. Median core volume for the training cohort was 32.4 mL which yields, per definition, a balanced class split. Applying this threshold to the independent test data resulted in 12 patients with large core volume and 11 patients with small core volume. Training duration for the 10-fold CV was in the range of 24 h on a local workstation (NVIDIA GeForce RTX 2070 Super) with online data augmentation and batch-wise data standardization. Evaluation and prediction of unseen data were of the order of a few seconds.

Figure 3 shows the ROC curves for the test folds within the 10-fold CV. The mean ROC curve is overlaid in blue, an interval of ± 1 SD is shaded in grey, and the dashed line represents random guessing. The mean (SD) ROC-AUC over 10 folds was 0.72 (0.10) for the test folds. In comparison, mean (SD) ROC-AUC for the validation folds, which were used to early stop training, was 0.75 (0.11). The averaged ensemble ROC-AUC for the independent test cohort selected from the ISLES 2018 challenge was 0.61, which is close to the ± 1 SD interval of the test folds from the training data. All results are summarized in Table 1.

The ablation study showed a decrease in mean (SD) ROC-AUC values for the reduced models, summarized in Table 2. Using the global feature extractor alone resulted in a ROC-AUC of 0.63 (0.14) and using the local feature extractor alone resulted in a ROC-AUC of 0.65 (0.13), compared to the full model with ROC-AUC 0.72 (0.10).

Table 1. Mean (SD) ROC-AUC of the final model for validation and test folds during CV and for the external test cohort. SD = standard deviation, CV = cross-validation, ROC-AUC = area under the receiver operator characteristics curve.

Validation Folds	Test Folds	Independent Test Cohort
0.75 (0.11)	0.72 (0.10)	0.61

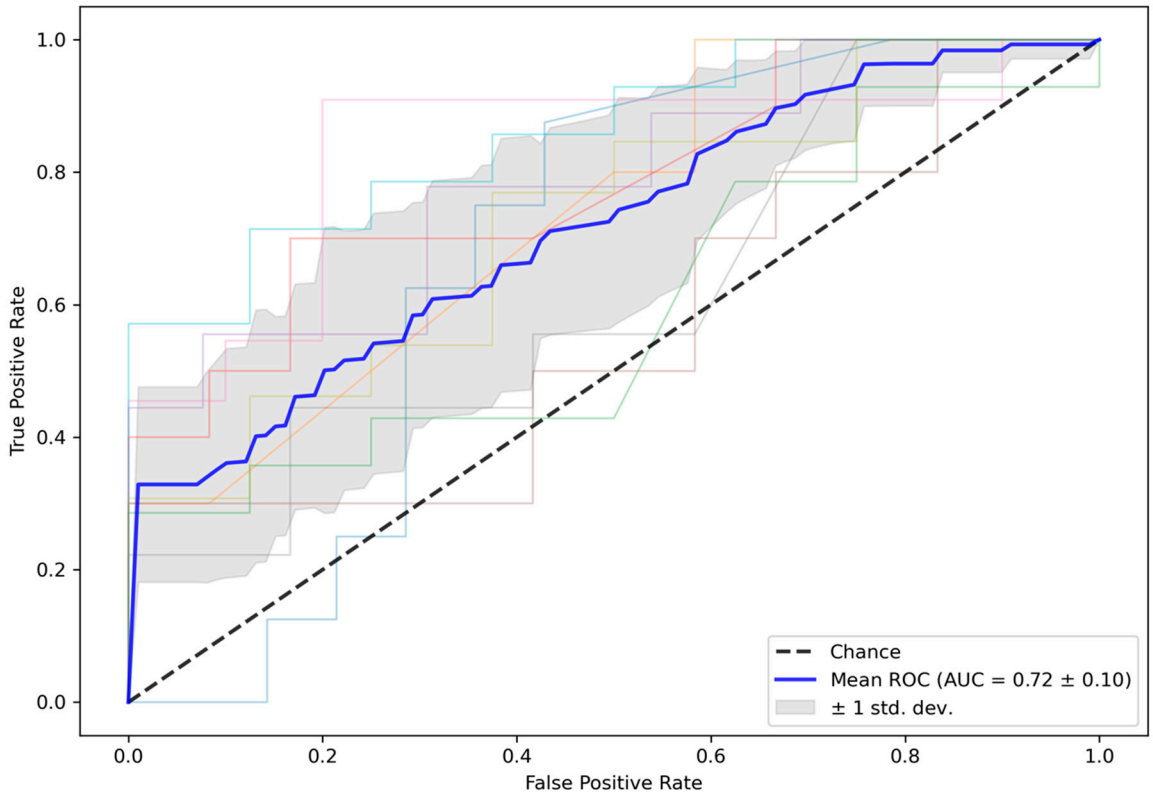

Figure 3. ROC curves for test data in the 10-fold CV. Mean (SD) ROC-AUC for the CV test folds was 0.72 (0.10). CV = cross-validation, ROC-AUC = area under the receiver operator characteristics curve.

Table 2. Mean (SD) ROC-AUC for the test folds during CV of the full model and the reduced models within the ablation study setting. SD = standard deviation, CV = cross-validation, ROC-AUC = area under the receiver operator characteristics curve.

Full Model	Global Feature Extractor Alone	Local Feature Extractor Alone
0.72 (0.10)	0.63 (0.14)	0.65 (0.13)

4. Discussion

In this proof-of-concept study, we developed a novel end-to-end deep learning approach to bypass conventional perfusion analysis, which allows to directly categorize patients into small or large core from raw CTP data without tracer-kinetic assumptions. We demonstrated this approach on 217 patients with acute ischemic stroke by directly predicting dichotomized infarct core volume and showed that the model learned relevant spatial and temporal features purely from the data. The results of the ablation study demonstrate the advantage of combined local and global temporal feature extraction, as only the full model yields the best performance. This corroborates our understanding that both short-term effects (e.g., sharp peaks in concentration) and long-term effects (e.g., wash-out) add relevant information and must therefore both be considered in the model architecture. In this proof-of-concept approach, our model cannot be translated to clinical practice immediately, however, achieved good predictive performance on an inhouse dataset in a 10-fold CV approach and generalized to independent test data, showing the potential of end-to-end CT perfusion analysis.

The proposed deep learning model is based on a 2D approach that covers a reduced portion of the middle cerebral artery territory, represented by the ASPECTS regions. Arguably, a 3D approach would contain more relevant information, but whole-brain CT perfusion is not available in all primary stroke centers [34]. Using the 2D approach, we were able to include all possible data, especially the external test data, which consisted

of two separate stacks of slices instead of whole-brain perfusion. We believe that this 2D approach is sufficient to prove the concept that spatial and temporal information can be extracted from CTP data to predict dichotomized core volume using deep learning.

In comparison to existing studies using deep learning for perfusion analysis in stroke CT [8], we focused on using the raw perfusion data solely. In contrast to voxelwise prediction of infarct status, the proposed model learned to predict dichotomized infarction core volume without taking additional parameters into account. While additional parameters like treatment information may be beneficial, additional user-provided information, such as a manually selected arterial input function, requires input from expert users and introduces user dependency. Our proposed model learned the link between perfusion input and tissue response purely based on the data and is free from tracer-kinetic assumptions.

Imaging-derived parameters play a crucial role in clinical decision-making in the setting of acute ischemic stroke. Foremost, CTP-derived ischemic core volume has become one of the key parameters in the decision for mechanical thrombectomy in the extended time window [35,36]. Currently, there is no consensus on the use of CTP-parameters for core/penumbra estimation. While software in large clinical trials used relative CBF thresholds (RAPID), other software relies on MTT (Philips Brain CT perfusion) or, in our case, on CBV. As relevant differences have been shown between vendors, our approach needs further validation for other CTP analysis thresholds [37]. In the present proof-of-concept study, we predicted dichotomized ischemic core volume as a simplified endpoint. Given sufficient training data, this approach can easily be generalized to more complex labels in future studies, such as impairment after discharge. The underlying relationship is harder to learn and may require incorporating additional clinical parameters into the model. Such deep-learning-based approaches may, therefore, be used to predict complications and even chronic functional outcomes in order to guide clinical management in and beyond the acute stroke phase.

The present study is not without limitations. First, the sample size of 217 training datasets is small for the complex problem of directly predicting an imaging-derived parameter from raw data and validation with a larger dataset is needed. To this end, dichotomized median core volume was chosen (i) to cast the problem as classification approach and (ii) to provide a balanced group distribution for the training dataset. This preliminary work may lay the groundwork for future studies to examine more clinically relevant endpoints, such as grade of disability or quality of life. Nevertheless, as proof of concept, our model achieved good results, which were validated on external test data. The performance gap between model predictions for in-house and external test data is likely due to differences in data quality and core labeling. Second, the number of perfusion timepoints of the CTP images was fixed for the model input, but appropriate interpolation could solve varying temporal resolutions.

5. Conclusions

In this proof-of-concept study, the proposed end-to-end deep learning approach bypasses conventional perfusion analysis and allows training a model that predicts dichotomized infarction core volume solely from slice-reduced CTP images without underlying tracer kinetic assumptions. Further studies can easily extend to additional clinically relevant endpoints.

Author Contributions: Conceptualization, all authors; methodology, A.M., B.S., P.W. and M.I.; data curation, A.M., P.R., M.P.F., L.K., S.T. and W.G.K.; writing—original draft preparation, A.M., P.R., M.P.F., B.S., P.W. and M.I.; writing—review and editing, A.M., P.R., M.P.F., B.S., P.W., B.E.-W., O.D., J.R., L.K., S.T., W.G.K. and M.I.; All authors have read and agreed to the published version of the manuscript.

Funding: A.M. was supported by the German Research Foundation (DFG) within the Research Training Group GRK 2274.

Institutional Review Board Statement: The retrospective study was approved by the institutional review board LMU Munich, 19-682, 2019-09-26, according to the Declaration of Helsinki of 2013.

Informed Consent Statement: The institutional review board LMU Munich waived the requirement for written informed consent (19-682, 2019-09-26) due to the retrospective analysis of anonymized data.

Data Availability Statement: Validation data available in a publicly accessible repository. The data of the ISLES 2018 challenge used as validation data in this study are openly available at https://www.smir.ch/ISLES/Start2018 (accessed on 2 May 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. van der Worp, H.B.; van Gijn, J. Acute Ischemic Stroke. *N. Engl. J. Med.* **2007**, *357*, 572–579. [CrossRef] [PubMed]
2. Chalet, L.; Boutelier, T.; Christen, T.; Raguenes, D.; Debatisse, J.; Eker, O.F.; Becker, G.; Nighoghossian, N.; Cho, T.-H.; Canet-Soulas, E.; et al. Clinical Imaging of the Penumbra in Ischemic Stroke: From the Concept to the Era of Mechanical Thrombectomy. *Front. Cardiovasc. Med.* **2022**, *9*, 438. [CrossRef] [PubMed]
3. Allmendinger, A.M.; Tang, E.R.; Lui, Y.W.; Spektor, V. Imaging of Stroke: Part 1, Perfusion CT–Overview of Imaging Technique, Interpretation Pearls, and Common Pitfalls. *AJR Am. J. Roentgenol.* **2012**, *198*, 52–62. [CrossRef] [PubMed]
4. Merino, J.G.; Warach, S. Imaging of Acute Stroke. *Nat. Rev. Neurol.* **2010**, *6*, 560–571. [CrossRef]
5. Ingrisch, M.; Sourbron, S. Tracer-Kinetic Modeling of Dynamic Contrast-Enhanced MRI and CT: A Primer. *J. Pharm. Pharm.* **2013**, *40*, 281–300. [CrossRef]
6. Sourbron, S.; Dujardin, M.; Makkat, S.; Luypaert, R. Pixel-by-Pixel Deconvolution of Bolus-Tracking Data: Optimization and Implementation. *Phys. Med. Biol.* **2007**, *52*, 429–447. [CrossRef]
7. Fieselmann, A.; Kowarschik, M.; Ganguly, A.; Hornegger, J.; Fahrig, R. Deconvolution-Based CT and MR Brain Perfusion Measurement: Theoretical Model Revisited and Practical Implementation Details. *J. Biomed. Imaging* **2011**, *2011*, 1–20. [CrossRef]
8. Bivard, A.; Levi, C.; Spratt, N.; Parsons, M. Perfusion CT in Acute Stroke: A Comprehensive Analysis of Infarct and Penumbra. *Radiology* **2013**, *267*, 543–550. [CrossRef]
9. Lansberg, M.G.; Christensen, S.; Kemp, S.; Mlynash, M.; Mishra, N.; Federau, C.; Tsai, J.P.; Kim, S.; Nogueria, R.G.; Jovin, T.; et al. Computed Tomographic Perfusion to Predict Response to Recanalization in Ischemic Stroke. *Ann. Neurol.* **2017**, *81*, 849–856. [CrossRef]
10. Alexandre, A.M.; Pedicelli, A.; Valente, I.; Scarcia, L.; Giubbolini, F.; D'Argento, F.; Lozupone, E.; Distefano, M.; Pilato, F.; Colosimo, C. May Endovascular Thrombectomy without CT Perfusion Improve Clinical Outcome? *Clin. Neurol. Neurosurg.* **2020**, *198*, 106207. [CrossRef]
11. Van Cauwenberge, M.G.A.; Dekeyzer, S.; Nikoubashman, O.; Dafotakis, M.; Wiesmann, M. Can Perfusion CT Unmask Postictal Stroke Mimics? A Case-Control Study of 133 Patients. *Neurology* **2018**, *91*, e1918–e1927. [CrossRef]
12. Becks, M.J.; Manniesing, R.; Vister, J.; Pegge, S.A.H.; Steens, S.C.A.; van Dijk, E.J.; Prokop, M.; Meijer, F.J.A. Brain CT Perfusion Improves Intracranial Vessel Occlusion Detection on CT Angiography. *J. Neuroradiol.* **2019**, *46*, 124–129. [CrossRef] [PubMed]
13. Wintermark, M.; Luby, M.; Bornstein, N.M.; Demchuk, A.; Fiehler, J.; Kudo, K.; Lees, K.R.; Liebeskind, D.S.; Michel, P.; Nogueira, R.G.; et al. International Survey of Acute Stroke Imaging Used to Make Revascularization Treatment Decisions. *Int. J. Stroke* **2015**, *10*, 759–762. [CrossRef] [PubMed]
14. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
15. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
16. Szegedy, C.; Toshev, A.; Erhan, D. Deep Neural Networks for Object Detection. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–10 December 2013; Burges, C.J.C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2013; Volume 26.
17. Yang, G.; Yu, S.; Dong, H.; Slabaugh, G.; Dragotti, P.L.; Ye, X.; Liu, F.; Arridge, S.; Keegan, J.; Guo, Y.; et al. DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. *IEEE Trans. Med. Imaging* **2018**, *37*, 1310–1321. [CrossRef]
18. Shen, D.; Wu, G.; Suk, H.-I. Deep Learning in Medical Image Analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [CrossRef]
19. Lai, M. Deep Learning for Medical Image Segmentation. *arXiv* **2015**, arXiv:1505.02000.
20. Ho, K.C.; Scalzo, F.; Sarma, K.V.; El-Saden, S.; Arnold, C.W. A Temporal Deep Learning Approach for MR Perfusion Parameter Estimation in Stroke. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 1315–1320.
21. Ulas, C.; Das, D.; Thrippleton, M.J.; Valdés Hernández, M.D.C.; Armitage, P.A.; Makin, S.D.; Wardlaw, J.M.; Menze, B.H. Convolutional Neural Networks for Direct Inference of Pharmacokinetic Parameters: Application to Stroke Dynamic Contrast-Enhanced MRI. *Front. Neurol.* **2019**, *9*, 1147. [CrossRef]

22. Robben, D.; Boers, A.M.M.; Marquering, H.A.; Langezaal, L.L.C.M.; Roos, Y.B.W.E.M.; van Oostenbrugge, R.J.; van Zwam, W.H.; Dippel, D.W.J.; Majoie, C.B.L.M.; van der Lugt, A.; et al. Prediction of Final Infarct Volume from Native CT Perfusion and Treatment Parameters Using Deep Learning. *Med. Image Anal.* **2020**, *59*, 101589. [CrossRef]
23. Chartrand, G.; Cheng, P.M.; Vorontsov, E.; Drozdzal, M.; Turcotte, S.; Pal, C.J.; Kadoury, S.; Tang, A. Deep Learning: A Primer for Radiologists. *RadioGraphics* **2017**, *37*, 2113–2131. [CrossRef] [PubMed]
24. Hakimelahi, R.; Yoo, A.J.; He, J.; Schwamm, L.H.; Lev, M.H.; Schaefer, P.W.; González, R.G. Rapid Identification of a Major Diffusion/Perfusion Mismatch in Distal Internal Carotid Artery or Middle Cerebral Artery Ischemic Stroke. *BMC Neurol.* **2012**, *12*, 132. [CrossRef] [PubMed]
25. Kistler, M.; Bonaretti, S.; Pfahrer, M.; Niklaus, R.; Büchler, P. The Virtual Skeleton Database: An Open Access Repository for Biomedical Research and Collaboration. *J. Med. Internet Res.* **2013**, *15*, e2930. [CrossRef] [PubMed]
26. Maier, O.; Menze, B.H.; von der Gablentz, J.; Häni, L.; Heinrich, M.P.; Liebrand, M.; Winzeck, S.; Basit, A.; Bentley, P.; Chen, L.; et al. ISLES 2015—A Public Evaluation Benchmark for Ischemic Stroke Lesion Segmentation from Multispectral MRI. *Med. Image Anal.* **2017**, *35*, 250–269. [CrossRef]
27. Barber, P.A.; Demchuk, A.M.; Zhang, J.; Buchan, A.M. Validity and Reliability of a Quantitative Computed Tomography Score in Predicting Outcome of Hyperacute Stroke before Thrombolytic Therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score. *Lancet* **2000**, *355*, 1670–1674. [CrossRef]
28. Van Rossum, G.; Drake, F.L., Jr. *Python Reference Manual*; Centrum voor Wiskunde en Informatica Amsterdam: Amsterdam, The Netherlands, 1995.
29. Beare, R.; Lowekamp, B.; Yaniv, Z. Image Segmentation, Registration and Characterization in R with SimpleITK. *J. Stat. Softw.* **2018**, *86*, 8. [CrossRef]
30. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
31. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. *arXiv* **2016**, arXiv:1603.04467.
32. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2015**, arXiv:1409.1556.
33. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Li, F.-F. Imagenet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
34. Almekhlafi, M.A.; Kunz, W.G.; Menon, B.K.; McTaggart, R.A.; Jayaraman, M.V.; Baxter, B.W.; Heck, D.; Frei, D.; Derdeyn, C.P.; Takagi, T.; et al. Imaging of Patients with Suspected Large-Vessel Occlusion at Primary Stroke Centers: Available Modalities and a Suggested Approach. *Am. J. Neuroradiol.* **2019**, *40*, 396–400. [CrossRef]
35. Albers, G.W.; Marks, M.P.; Kemp, S.; Christensen, S.; Tsai, J.P.; Ortega-Gutierrez, S.; McTaggart, R.A.; Torbey, M.T.; Kim-Tenser, M.; Leslie-Mazwi, T.; et al. Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging. *N. Engl. J. Med.* **2018**, *378*, 708–718. [CrossRef] [PubMed]
36. Nogueira, R.G.; Jadhav, A.P.; Haussen, D.C.; Bonafe, A.; Budzik, R.F.; Bhuva, P.; Yavagal, D.R.; Ribo, M.; Cognard, C.; Hanel, R.A.; et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. *N. Engl. J. Med.* **2018**, *378*, 11–21. [CrossRef] [PubMed]
37. Austein, F.; Riedel, C.; Kerby, T.; Meyne, J.; Binder, A.; Lindner, T.; Huhndorf, M.; Wodarg, F.; Jansen, O. Comparison of Perfusion CT Software to Predict the Final Infarct Volume After Thrombectomy. *Stroke* **2016**, *47*, 2311–2317. [CrossRef] [PubMed]

6 | Conclusion

This works demonstrates a robust evaluation of dynamic contrast-enhanced (DCE) imaging data for perfusion quantification by means of two novel analysis methods. On the one hand, reliable quantification of perfusion parameters is addressed by a Bayesian formulation of tracer-kinetic modeling which yields parameter distributions rather than point estimates. This allows the uncertainty of the parameters to be included in the analysis which increases the reliability of the estimates and therefore clinical decision making.

On the other hand, approximations in the context of tracer-kinetic modeling are circumvented by the use of convolutional neural networks (CNNs). CNNs allow to learn spatial and temporal features from the raw imaging data to directly predict clinical endpoints and therefore completely bypass tracer-kinetic modeling. Building on the foundations of the present work, future research could even combine both approaches in Bayesian neural networks. This is an emerging research area and promises to increase the robustness of perfusion quantification and therefore clinical acceptance even further.

Bibliography

[1] D. L. Thomas, M. F. Lythgoe, G. S. Pell, F. Calamante, and R. J. Ordidge, "The measurement of diffusion and perfusion in biological systems using magnetic resonance imaging", *Physics in Medicine and Biology*, vol. 45, no. 8, R97–R138, Jul. 2000, Publisher: IOP Publishing, ISSN: 0031-9155. DOI: 10.1088/0031-9155/45/8/201.

[2] H. B. van der Worp and J. van Gijn, "Acute Ischemic Stroke", *New England Journal of Medicine*, vol. 357, no. 6, pp. 572–579, Aug. 2007, Publisher: Massachusetts Medical Society, ISSN: 0028-4793. DOI: 10.1056/NEJMcp072057.

[3] J. L. Anderson and D. A. Morrow, "Acute Myocardial Infarction", *New England Journal of Medicine*, vol. 376, no. 21, pp. 2053–2064, May 2017, ISSN: 0028-4793. DOI: 10.1056/NEJMra1606915.

[4] R. J. Gillies, P. A. Schomack, T. W. Secomb, and N. Raghunand, "Causes and Effects of Heterogeneous Perfusion in Tumors", *Neoplasia*, vol. 1, no. 3, pp. 197–207, Aug. 1999, ISSN: 1476-5586. DOI: 10.1038/sj.neo.7900037.

[5] J. P. B. O'Connor, A. Jackson, G. J. M. Parker, and G. C. Jayson, "DCE-MRI biomarkers in the clinical evaluation of antiangiogenic and vascular disrupting agents", *British Journal of Cancer*, vol. 96, no. 2, pp. 189–195, Jan. 2007, ISSN: 0007-0920. DOI: 10.1038/sj.bjc.6603515.

[6] A. Jackson, J. P. O'Connor, G. J. Parker, and G. C. Jayson, "Imaging Tumor Vascular Heterogeneity and Angiogenesis using Dynamic Contrast-Enhanced Magnetic Resonance Imaging", *Clinical Cancer Research*, vol. 13, no. 12, pp. 3449–3459, Jun. 2007, ISSN: 1078-0432. DOI: 10.1158/1078-0432.CCR-07-0238.

[7] R. Bammer, "Basic principles of diffusion-weighted imaging", *European Journal of Radiology*, vol. 45, no. 3, pp. 169–184, Mar. 2003, ISSN: 0720-048X. DOI: 10.1016/S0720-048X(02)00303-0.

[8] R. B. Buxton, L. R. Frank, E. C. Wong, B. Siewert, S. Warach, and R. R. Edelman, "A general kinetic model for quantitative perfusion imaging with arterial spin labeling", *Magnetic Resonance in Medicine*, vol. 40, no. 3, pp. 383–396, 1998, ISSN: 1522-2594. DOI: 10.1002/mrm.1910400308.

[9] J. P. B. O'Connor, P. S. Tofts, K. A. Miles, L. M. Parkes, G. Thompson, and A. Jackson, "Dynamic contrast-enhanced imaging techniques: CT and MRI", *The*

British Journal of Radiology, vol. 84, no. special_issue_2, S112–S120, Dec. 2011, ISSN: 0007-1285. DOI: 10.1259/bjr/55166688.

[10] M. Laniado, H. J. Weinmann, W. Schörner, R. Felix, and U. Speck, "First use of GdDTPA/dimeglumine in man", *Physiological chemistry and physics and medical NMR*, vol. 16, no. 2, pp. 157–165, 1984, ISSN: 0748-6642.

[11] J. Lohrke, T. Frenzel, J. Endrikat, F. C. Alves, T. M. Grist, M. Law, J. M. Lee, T. Leiner, K.-C. Li, K. Nikolaou, M. R. Prince, H. H. Schild, J. C. Weinreb, K. Yoshikawa, and H. Pietsch, "25 Years of Contrast-Enhanced MRI: Developments, Current Challenges and Future Perspectives", *Advances in Therapy*, vol. 33, no. 1, pp. 1–28, Jan. 2016, ISSN: 1865-8652. DOI: 10.1007/s12325-015-0275-4.

[12] C. Roberts, B. Issa, A. Stone, A. Jackson, J. C. Waterton, and G. J. M. Parker, "Comparative study into the robustness of compartmental modeling and model-free analysis in DCE-MRI studies", *Journal of magnetic resonance imaging: JMRI*, vol. 23, no. 4, pp. 554–563, Apr. 2006, ISSN: 1053-1807. DOI: 10.1002/jmri.20529.

[13] S. P. Sourbron and D. L. Buckley, "Tracer kinetic modelling in MRI: Estimating perfusion and capillary permeability", *Physics in Medicine and Biology*, vol. 57, no. 2, R1–R33, Dec. 2011, ISSN: 0031-9155. DOI: 10.1088/0031-9155/57/2/R1.

[14] T. S. Koh, S. Bisdas, D. M. Koh, and C. H. Thng, "Fundamentals of tracer kinetics for dynamic contrast-enhanced MRI", *Journal of Magnetic Resonance Imaging*, vol. 34, no. 6, pp. 1262–1276, 2011, ISSN: 1522-2586. DOI: 10.1002/jmri.22795.

[15] S. P. Sourbron and D. L. Buckley, "Classic models for dynamic contrast-enhanced MRI", *NMR in Biomedicine*, vol. 26, no. 8, pp. 1004–1027, 2013, ISSN: 1099-1492. DOI: 10.1002/nbm.2940.

[16] M. Ingrisch and S. Sourbron, "Tracer-kinetic modeling of dynamic contrast-enhanced MRI and CT: A primer", *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 40, no. 3, pp. 281–300, Jun. 2013, ISSN: 1573-8744. DOI: 10.1007/s10928-013-9315-3.

[17] L. Beuzit, P.-A. Eliat, V. Brun, J.-C. Ferré, Y. Gandon, E. Bannier, and H. Saint-Jalmes, "Dynamic contrast-enhanced MRI: Study of inter-software accuracy and reproducibility using simulated and clinical data", *Journal of magnetic resonance imaging: JMRI*, vol. 43, no. 6, pp. 1288–1300, 2016, ISSN: 1522-2586. DOI: 10.1002/jmri.25101.

[18] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: physical principles and sequence design*, 2nd edition. Hoboken, New Jersey: John Wiley & Sons, Inc, 2014, ISBN: 978-1-118-63395-3.

[19] M. H. Levitt, *Spin dynamics: basics of nuclear magnetic resonance*, 2nd ed. Chichester, England ; Hoboken, NJ: John Wiley & Sons, 2008, ISBN: 978-0-470-51117-6.

[20] M. Ingrisch, "Quantification of cerebral hemodynamics with dynamic contrast-enhanced MRI", Text.PhDThesis, Ludwig-Maximilians-Universität München, Oct. 2012.

[21] W. A. Kalender, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*. John Wiley & Sons, Jul. 2011, ISBN: 978-3-89578-644-0.

[22] T. M. Buzug, *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Berlin Heidelberg: Springer-Verlag, 2008, ISBN: 978-3-642-07257-4.

[23] G. Poludniowski, A. Omar, R. Bujila, and P. Andreo, "Technical Note: SpekPy v2.0—a software toolkit for modeling x-ray tube spectra", *Medical Physics*, vol. 48, no. 7, pp. 3630–3637, 2021, ISSN: 2473-4209. DOI: 10.1002/mp.14945.

[24] S. Seltzer, *XCOM-Photon Cross Sections Database, NIST Standard Reference Database 8*, 1987.

[25] S. Sourbron, "Technical aspects of MR perfusion", *European Journal of Radiology*, Diffusion and Perfusion of the Body, vol. 76, no. 3, pp. 304–313, Dec. 2010, ISSN: 0720-048X. DOI: 10.1016/j.ejrad.2010.02.017.

[26] F. Calamante, "Arterial input function in perfusion MRI: A comprehensive review", *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 74, pp. 1–32, Oct. 2013, ISSN: 1873-3301. DOI: 10.1016/j.pnmrs.2013.04.002.

[27] A. E. Hansen, H. Pedersen, E. Rostrup, and H. B. Larsson, "Partial volume effect (PVE) on the arterial input function (AIF) in T1-weighted perfusion imaging and limitations of the multiplicative rescaling approach", *Magnetic Resonance in Medicine*, vol. 62, no. 4, pp. 1055–1059, 2009, ISSN: 1522-2594. DOI: 10.1002/mrm.22098.

[28] G. J. Parker, C. Roberts, A. Macdonald, G. A. Buonaccorsi, S. Cheung, D. L. Buckley, A. Jackson, Y. Watson, K. Davies, and G. C. Jayson, "Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI", *Magnetic Resonance in Medicine*, vol. 56, no. 5, pp. 993–1000, 2006, ISSN: 1522-2594. DOI: 10.1002/mrm.21066.

[29] K. A. Miles, M. R. Griffiths, and M. A. Fuentes, "Standardized Perfusion Value: Universal CT Contrast Enhancement Scale that Correlates with FDG PET in Lung Nodules", *Radiology*, vol. 220, no. 2, pp. 548–553, Aug. 2001, ISSN: 0033-8419. DOI: 10.1148/radiology.220.2.r01au26548.

[30] J. Pintaske, P. Martirosian, H. Graf, G. Erb, K.-P. Lodemann, C. D. Claussen, and F. Schick, "Relaxivity of Gadopentetate Dimeglumine (Magnevist), Gadobutrol (Gadovist), and Gadobenate Dimeglumine (MultiHance) in human blood plasma at 0.2, 1.5, and 3 Tesla", *Investigative Radiology*, vol. 41, no. 3, pp. 213–221, Mar. 2006, ISSN: 0020-9996. DOI: 10.1097/01.rli.0000197668.44926.f7.

[31] S. Sourbron, M. Dujardin, S. Makkat, and R. Luypaert, "Pixel-by-pixel deconvolution of bolus-tracking data: Optimization and implementation", *Physics in Medicine and Biology*, vol. 52, no. 2, pp. 429–447, Jan. 2007, ISSN: 0031-9155. DOI: 10.1088/0031-9155/52/2/009.

[32] G. Brix, F. Kiessling, R. Lucht, S. Darai, K. Wasser, S. Delorme, and J. Griebel, "Microcirculation and microvasculature in breast tumors: Pharmacokinetic analysis of dynamic MR image series", *Magnetic Resonance in Medicine*, vol. 52, no. 2, pp. 420–429, 2004, ISSN: 1522-2594. DOI: 10.1002/mrm.20161.

[33] P. S. Tofts and A. G. Kermode, "Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts", *Magnetic Resonance in Medicine*, vol. 17, no. 2, pp. 357–367, Feb. 1991, ISSN: 0740-3194.

[34] P. S. Tofts, "Modeling tracer kinetics in dynamic Gd-DTPA MR imaging", *Journal of magnetic resonance imaging: JMRI*, vol. 7, no. 1, pp. 91–101, Feb. 1997, ISSN: 1053-1807.

[35] S. P. Sourbron and D. L. Buckley, "On the scope and interpretation of the Tofts models for DCE-MRI", *Magnetic Resonance in Medicine*, vol. 66, no. 3, pp. 735–745, Sep. 2011, ISSN: 1522-2594. DOI: 10.1002/mrm.22861.

[36] P. S. Tofts, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp, H. B. Larsson, T. Y. Lee, N. A. Mayr, G. J. Parker, R. E. Port, J. Taylor, and R. M. Weisskoff, "Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: Standardized quantities and symbols", *Journal of magnetic resonance imaging: JMRI*, vol. 10, no. 3, pp. 223–232, Sep. 1999, ISSN: 1053-1807.

[37] G. A. F. Seber and C. J. Wild, *Nonlinear Regression*. John Wiley & Sons, Sep. 2003, ISBN: 978-0-471-47135-6.

[38] R. McElreath, *Statistical rethinking: a Bayesian course with examples in R and Stan*, 1st. Chapman and Hall/CRC, 2015, ISBN: 978-1-4822-5344-3.

[39] J. VanderPlas, *Frequentism and Bayesianism: A Python-driven Primer*, Nov. 2014.

[40] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis, Third Edition*, 3 edition. Boca Raton, Fla.: Chapman and Hall/CRC, Nov. 2013, ISBN: 978-1-4398-4095-5.

[41] W. R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. CRC Press, Dec. 1995, ISBN: 978-0-412-05551-5.

[42] J. Kruschke, *Doing Bayesian Data Analysis*, 2nd. Academic Press, 2014, ISBN: 978-0-12-405888-0.

[43] M. D. Hoffman and A. Gelman, "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo", *ArXiv e-prints*, vol. 1111, arXiv:1111.4246, Nov. 2011.

[44] V. J. Schmid, B. Whitcher, A. R. Padhani, N. J. Taylor, and G.-Z. Yang, "Bayesian methods for pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging", *IEEE transactions on medical imaging*, vol. 25, no. 12, pp. 1627–1636, Dec. 2006, ISSN: 0278-0062. DOI: 10.1109/TMI.2006.884210.

[45] M. R. Orton, D. J. Collins, S. Walker-Samuel, J. A. d'Arcy, D. J. Hawkes, D. Atkinson, and M. O. Leach, "Bayesian estimation of pharmacokinetic pa-

rameters for DCE-MRI with a robust treatment of enhancement onset time", *Physics in Medicine & Biology*, vol. 52, no. 9, p. 2393, 2007, ISSN: 0031-9155. DOI: 10.1088/0031-9155/52/9/005.

[46] M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. M. Smith, "Bayesian analysis of neuroimaging data in FSL", *NeuroImage*, vol. 45, no. 1 Suppl, S173–186, Mar. 2009, ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2008.10.055.

[47] A. Tietze, A. Nielsen, I. K. Mikkelsen, M. B. Hansen, A. Obel, L. Østergaard, and K. Mouridsen, "Bayesian modeling of Dynamic Contrast Enhanced MRI data in cerebral glioma patients improves the diagnostic quality of hemodynamic parameter maps", *PLOS ONE*, vol. 13, no. 9, e0202906, Sep. 2018, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0202906.

[48] A. Shukla-Dave, N. A. Obuchowski, T. L. Chenevert, S. Jambawalikar, L. H. Schwartz, D. Malyarenko, W. Huang, S. M. Noworolski, R. J. Young, M. S. Shiroishi, H. Kim, C. Coolens, H. Laue, C. Chung, M. Rosen, M. Boss, and E. F. Jackson, "Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials", *Journal of Magnetic Resonance Imaging*, vol. 49, no. 7, e101–e121, Nov. 2018, ISSN: 1522-2586. DOI: 10.1002/jmri.26518.

[49] M. A. Nielsen, "Neural Networks and Deep Learning", 2015.

[50] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[51] P. M. Cheng, E. Montagnon, R. Yamashita, I. Pan, A. Cadrin-Chênevert, F. Perdigón Romero, G. Chartrand, S. Kadoury, and A. Tang, "Deep Learning: An Update for Radiologists", *RadioGraphics*, vol. 41, no. 5, pp. 1427–1445, Sep. 2021, Publisher: Radiological Society of North America, ISSN: 0271-5333. DOI: 10.1148/rg.2021200210.

[52] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A Probabilistic Programming Language", *Journal of Statistical Software*, vol. 76, no. 1, pp. 1–32, Jan. 2017, ISSN: 1548-7660. DOI: 10.18637/jss.v076.i01.

[53] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems", 2015, Software available from tensorflow.org.

List of Figures

Danksagung

Michi, dir mech i als ersts danke song, ohne di warad i wirklich ned so weit kema und es häd a nur hoib sovui Spaß gmacht! Danke für de Beste Betreuung de ma si überhapt nur vorstoin ko, für ois wosd ma zoagd hosd in de mittlerweile 5 Joa wo i scho bei dir bin. Es hod ma narrisch guad gfoin und i find vo dir kinnan si alle a Scheibal obschnein - Merci!

Birgit, danke a an di für dei Betreuung und dass dus möglich gmacht hosd, dass i an Auslandsaufenthalt bei dir in Toronto macha hob kinna. Des war glabi oane vo de besten Erfahrungen de i gmacht hob!

Olaf, dir gebührt genauso mei Dank für dei Betreuung. I hob gwusst i konn immer zu dir kema mit Fragen - de Antwort war zwar ned immer des wos i hern woid, aber wissenschaftlich wars immer ein Gewinn!

Danke an des ganze Clinical Data Science Team: Balthi, Philipp, Tessi, Moritz, es is a Traum mit eich zu diskutieren, sich neie Ideen zum überleng und ab und zua amoi a Rundn Kartn zum spuin! Ohne eich wars ganz sche fad gwen moan i! In diesem Sinne, wer gibt?

Danke a an alle im Josef-Lissner-Labor für de ganzen netten Mittagessen, Diskussionsrunden und Kaffeepausen!

Am End mech i mi no bedanken bei meine Freind, meiner Family und meiner Freindin. So a Abat schreibt si ned immer leicht und ihr warts ollawei für mi do wenns gfeid hod. Meistens hods aber eh basd und bei freudigen Ereignissen koma sowieso auf eich zoin - Merci!