COPENHAGEN UNIVERSITY HOSPITAL, RIGSHOSPITALET
NEUROBIOLOGY RESEARCH UNIT

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Comparison of prospective and retrospective motion correction for Magnetic Resonance Imaging of the brain

Hannah Susanne Eichhorn

Master's Thesis in Physics

May 20th 2021

Supervisors:
Asst. Prof. Melanie Ganz-Benjaminsen, PhD
Assoc. Prof. Heloisa Nunes Bordallo, PhD

Abstract

Head motion is one of the most common sources of artefacts for Magnetic Resonance Imaging (MRI) of the brain. Especially children, being intimidated by the dimensions and the noise of the scanner, tend to move considerably during image acquisition. Thus, the use of general anaesthesia or sedation is common practice in clinical routines to avoid motion artefacts. However, general anaesthesia requires additional equipment and personnel and concerns about potential drug-related risks are increasing. With this thesis, I aim to compare different methods for motion correction of MRI as alternative strategies for avoiding motion artefacts during clinical MRI examinations. I evaluate the performance of prospective and retrospective motion correction on a set of six 2D- and 3D-encoded sequences from a clinical pediatric MR protocol. For each sequence, we acquired scans with and without a predefined motion pattern in 22 healthy volunteers. I analyse and statistically compare a set of image quality measures as well as observer quality scores. Furthermore, I quantify the influence of motion correction on motion-related changes in cortical thickness estimates using a general linear model.

In this work, I show that for the 3D-encoded sequences, image quality considerably increases from scans without motion correction, over retrospectively to prospectively corrected scans. Selectively reacquiring slices with the highest level of motion additionally improved image quality. Furthermore, cortical thickness estimates from motion corrected scans do not correlate with motion to the same extent as observed for uncorrected scans. For three of the 2D-encoded sequences I observe higher image quality due motion correction and reacquisition compared to the clinical standard, even though the effect size varies between the different sequences. I did not observe an improvement due to motion correction in the T_2* sequence, which is known to be highly sensitive to field inhomogeneities. However, this sequence can be substituted by an accelerated, easily repeatable sequence.

My findings confirm that the use of prospective motion correction is feasible for a full clinical children's protocol, reducing motion artefacts and ensuring diagnostic image quality. As a next step in establishing motion correction in clinical routines, the developed and validated set up will be tested on a clinical pediatric population.

Preface

This thesis was prepared at the Neurobiology Research Unit, Copenhagen University Hospital, Rigshospitalet in fulfillment of the requirements for acquiring a Master of Science degree in Physics, with a specialisation in Bio- and Medical Physics.

The work has been carried out from August 2020 to May 2021 and the assigned workload is 60 ECTS credits.

The thesis aims to compare prospective and retrospective motion correction techniques for a full clinical Magnetic Resonance Imaging protocol of the brain.

The thesis was handed in on May 20th 2021.

Supervisors
Asst. Prof. Melanie Ganz-Benjaminsen, PhD, Rigshospitalet, Department of Computer Science, University of Copenhagen
Assoc. Prof. Heloisa Nunes Bordallo, PhD, Niels-Bohr Institute, University of Copenhagen

Examiner
Assoc. Prof. Lars G. Hanson, PhD, Technical University of Denmark, Danish Research Centre For Magnetic Resonance

Copenhagen, May 20th 2021

Hannah Susanne Eichhorn

Acknowledgements

I would like to express my sincere gratitude to all those who supported me during completing this thesis and contributed to the MoCo volunteer study. In especially, I would like to thank

Asst. Prof. Melanie Ganz-Benjaminsen, the best supervisor I could possibly imagine, for her dedicated involvement throughout the project, her invaluable advise and for always having an open door for my questions.

Assoc. Prof. Heloisa Nunes Bordallo who kindly co-supervised my thesis, always offering important support and her valued experience, whenever I approached her for help.

The whole Neurobiology Research Unit at Rigshospitalet and in particularly NRU's data analysis group for welcoming me with open arms in their team. Special thanks go to Chief Engineer Claus Svarer and my office colleague Anders Stevnhoved Olsen for their valuable input throughout the thesis, as well as to the MoCo project nurse Asta Kongsgaard for her flexible recruiting and many nice scan sessions.

Assoc. Prof. Andre van der Kouwe and Instructor Robert Frost from the Athinoula A. Martinos Center for Biomedical Imaging (Massachusetts, US) for providing us with implementations of prospective motion correction and for offering many valuable discussions throughout the project.

Ph.D. student Jakob Slipsager from the Department of Applied Mathematics and Computer Science (Technical University of Denmark) who kindly helped with the retrospective correction of our dataset, as well as transferring data back and forth to TracInnovations.

MR physicist Karen Kettless (Siemens Healthineers, Denmark) for her patience, when supporting us in setting up the MR protocol, and her knowledge, helping me to understand sequence-related details.

Senior Researcher Kathrine Skak Madsen, as well as radiographer students Martin Riis Rassing and Bianca Pedersen who contributed to this work with discussions about image quality and with quality scoring of the MR images.

CTO Stefan Glimberg and Software Developer Claus Benjaminsen (TracInnovations, Denmark) for their support regarding questions connected to the motion tracking.

The Novo Scholarship Programme and in especially my mentor, Peter Helding Kvist, for his kind support and for offering a different point of view.

My good friends **Katharina Kohler and Alexandra Baranik**, as well as my dear sister **Leonie Eichhorn** for proof-reading this work and offering several helpful suggestions.

My family who has always greatly supported me, made a large number of things possible for me and was there for me at any time.

Nomenclature

ADC	Apparent Diffusion Coefficient
CSF	Cerebrospinal Fluid
DW	Diffusion Weighted
DWI	Diffusion Weighted Imaging
EPI	Echo Planar Imaging
FDR	False Discovery Rate
FID	Free Induction Decay
FLAIR	Fluid-Attenuated Inversion Recovery
FLASH	Fast Low Angle Shot
FOV	Field of View
GA	General anaesthesia
GRE	Gradient Echo
MoCo	Motion Correction
MPRAGE	Magnetization Prepared Rapid Gradient Echo
MRI	Magnetic Resonance Imaging
PE	Phase Encoding
pMC	prospective Motion Correction
PSNR	Peak Signal-to-Noise Ratio
RAS	Right-Anterior-Superior coordinate system
RF	Radio Frequency
rMC	retrospective Motion Correction
RMS	Root Mean Squared
SAR	Specific Absorption Rate
SPACE	Sampling Perfection with Application optimised Contracts using different flip angle Evolutions
SSIM	Structural Similarity Index
STIR	Short-TI Inversion Recovery
SWI	Susceptibility Weighted Imaging
TGRAD	Tenengrad
TSE	Turbo Spin Echo

Notation

The multiplication of a vector $\vec{v}$ with a scalar a is denoted by $a \cdot \vec{v}$. Analogously, the inner product of two vectors $\vec{v}$ and $\vec{w}$ is denoted with $\vec{v} \cdot \vec{w} = \sum_i v_i w_i$. In the case of two matrices A and B, $C = A \cdot B$ indicates matrix multiplication with $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

The cross product of two vectors $\vec{v}$ and $\vec{w}$ is denoted with $\vec{v} \times \vec{w} =\mid \vec{v} \mid\mid \vec{w} \mid \sin(\theta) \cdot \vec{n}$, with $\mid \vec{v} \mid$ representing the magnitude of the vector $\vec{v}$, θ the angle between the two vectors and $\vec{n}$ representing the unit vector perpendicular to the plane containing $\vec{v}$ and $\vec{w}$.

Contents

1 Introduction

Magnetic Resonance Imaging (MRI) is a medical imaging technique relying on the properties of the spins of nuclei in the presence of an external magnetic field. Advantages compared to other modalities such as Computed Tomography or Positron Emission Tomography, are its non-invasive nature and high contrast in soft tissue, i.e. in the brain [1]. However, due to long acquisition times of 30 to 60 minutes, motion is the most frequent cause of artefacts [2, 3]. Motion artefacts might render affected images unsuitable for diagnostics, impeding radiologic interpretations as well as the estimation of morphological measures as e.g. cortical thickness or gray matter volume [4, 5]. Andre et al. [6] report that approximately one fifth of all MRI examinations need to be partly repeated, leading to an estimated cost of 115,000 US dollars per scanner per year.

Especially children are frightened by the large size of the scanner, the limited space within the bore and the loud noise, which leads to significantly more motion of children during image acquisition compared to adults [7]. Until now, for most children between 4 and 10 years motion artefacts are avoided by preventing motion with deep sedation or general anaesthesia (GA) [8]. However, GA is associated with long waiting periods, a more complicated setting and higher costs [8, 9]. Moreover, concerns about potential risks related to drug interactions are increasing [2, 7, 10]. Havidich et al. [11] found an occurrence of adverse events (like airway obstruction, coughing or oxygen desaturation) of 8.6% in patients between 0 and 22 years who were anaesthetized / sedated for diagnostic or therapeutic procedures.

In the literature, different strategies are proposed for avoiding GA and at the same time ensuring high image quality, when young children need to undergo MRI. One approach is to prepare the children adequately before the examination, as well as to entertain or distract them during image acquisition. De Bie et al. [7] showed that GA could be avoided for 81 out of 90 children, when the children are trained with a mock scanner prior to the real MR scan. Similarly, Törnqvist et al. [3] reported acceptable image quality for 30 out of 33 children, after preparing them with a story book as well as with a small model of the MR scanner and offering distraction with a movie during acquisition.

Another way to improve image quality, even though the patient is moving, is to perform motion correction (MoCo) during or after image acquisition. Different techniques for MR motion correction exist, some of them using external motion tracking, others estimating motion from the image data itself. Both prospective, as well as retrospective approaches have been reported in the literature to avoid motion artefacts and increase overall image quality [12–15]. Nevertheless, in the literature motion correction has so far mostly been applied to 3D-encoded sequences and only a small number of motion correction applications are used in a clinical setting [12].

This difficulty is addressed in the MoCo project at the Neurobiology Research Unit, where we aim to show that high image quality can be realized for small children in a clinical setting without the use of GA, if they are trained with an app and a mock scanner beforehand and if motion correction is applied during image acquisition. Therefore, the performances of different motion correction techniques need to be tested for a full clinical protocol consisting of both 2D- and 3D-encoded sequences. Within this thesis, I compare prospective and retrospective motion correction on a dataset we acquired in 22 healthy volunteers as a first step of validating that motion correction improves image quality in a clinical setting, as well.

2 Theoretical Background

The aim of this chapter is to provide an overview of the underlying physical principles of Magnetic Resonance Imaging, the influence of motion on image formation as well as state-of-the-art motion correction techniques, primarily based on Godenschweger et al. [12] as well as textbooks by Brown et al. [1], Schlegel et al. [16] and Brix et al. [17].

2.1 Principles of Magnetic Resonance Imaging

2.1.1 Nuclear Spin and Macroscopic Magnetization

In atomic nuclei spin angular momentum and orbital angular momentum of bound protons and neutrons couple to a total angular momentum I, commonly referred to as nuclear spin [16]. Being fermions, individual protons and neutrons have spin $s = \frac{1}{2}$. In nuclei two spins generally combine and cancel out, separately. Consequently, in the ground state all nuclei with an odd number of protons or neutrons will have a non-zero nuclear spin.

The magnetic moment $\vec{\mu}$ of a nucleus is proportional to its nuclear spin $\vec{I}$:

$$\vec{\mu} = \gamma \cdot \vec{I}, \tag{1}$$

with γ representing the gyromagnetic ratio, a characteristic property of each nucleus. In the case of hydrogen: $\gamma = 26.752 \cdot 10^7 \ \frac{radMHz}{T}$. Within this thesis, only hydrogen MRI will be performed, which is why the following discussion focuses on hydrogen atoms, also referred to as protons.

For an MR experiment, all protons in a sample are taken into consideration. Thus, a net magnetization needs to be introduced. This macroscopic magnetization vector $\vec{M}(\vec{r}, t)$ is commonly defined as average over all N magnetic moments in a certain volume V, which are assumed to have the same resonance frequency:

$$\vec{M}(\vec{r}, t) = \frac{1}{V} \sum_{i=0}^{N} \vec{\mu}_i. \tag{2}$$

In the presence of an external magnetic field $\vec{B}_0$, the spins align parallel or anti-parallel to the magnetic field and the energy levels of ^{1}H split up into $2I + 1 = 2$ sub levels, corresponding to the different orientations of the magnetic moment relative to the magnetic field, which is known as Nuclear Zeeman Effect. The energy difference between the levels is given by $\omega_0 = \gamma B_0$, which is known as resonance frequency and determines the frequency of the Radio Frequency (RF) pulses which will be used in the following MR experiment to excite the spins.

At a given temperature T, the Boltzmann statistics define the number of magnetic moments for each quantum number at thermal equilibrium:

$$N_m = \frac{N}{Z} \exp(-\frac{E_m}{k_B T}), \tag{3}$$

with the Boltzmann's constant $k_B = 1.38 \cdot 10^{-23} \frac{J}{K}$ and the partition function Z. Using the high temperature approximation $E_m = \gamma m \hbar B_0 \ll k_B T$, equations (2) and (3) result in the following equilibrium magnetization:

$$M_0 \approx \frac{N}{V} \frac{\gamma^2 \hbar^2 B_0 I(I+1)}{3 k_B T}. \tag{4}$$

2.1.2 Precession and Excitation

An external magnetic field $\vec{B_{ext}}$ causes a torque on each single spin. Thus, the equation of motion for the macroscopic magnetization is given by summing up all individual equations, leading to:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B_{ext}}. \tag{5}$$

If the macroscopic magnetization is not oriented parallel to the magnetic field, the solution of this equation for a constant $\vec{B_{ext}} = B_0\vec{e_z}$ will be a precession around the z-axis with the Larmor frequency

$$\omega_L = \gamma B_0, \tag{6}$$

which matches the resonance frequency ω_0 introduced in section 2.1.1. This precession of the spins induces the signal measured by the receiver coil of an MR experiment.

However, at equilibrium the spins are aligned with the magnetic field and need to be tipped to an excited state for detecting their signal. This is realised with a RF pulse, causing an additional, time-dependent magnetic field, $\vec{B_1}$, via transmit coils:

$$\vec{B_1}(t) = B_1(\cos(\omega_{RF}t) \cdot \vec{e_x} + \sin(\omega_{RF}t) \cdot \vec{e_y}). \tag{7}$$

To simplify solving equation (5) with this magnetic field, a rotating reference frame is introduced in which the $\vec{B_1}$-field is at rest [1]. The coordinates in this frame will be primed. The effective total field in the rotating frame is given by the following equation, in which $\vec{B_1}$ is arbitrarily chosen in x'-direction:

$$\vec{B_{eff}}(t) = (B_0 - \frac{\omega_{RF}}{\gamma}) \cdot \vec{e_z} + B_1 \cdot \vec{e_{x'}}. \tag{8}$$

If the frequency of the RF pulse is equivalent to the resonance frequency ω_0, the macroscopic magnetization will only be influenced by a static B_1-field. Consequently, for $\vec{B_1}$ chosen in x'-direction $\vec{M}$ will precess around the x'-axis with the frequency $\omega_1 = \gamma B_1$. In the laboratory frame this corresponds to a helical motion of $\vec{M}$ on the surface of a sphere around B_0.

If the macroscopic magnetization was in the equilibrium state previous to the RF pulse of duration t_p, the total angle by which the magnetization will be tipped, commonly called flip angle, can be calculated with

$$\alpha = \int_0^{t_p} \gamma B_1(t)dt. \tag{9}$$

2.1.3 Relaxation

For the above discussed phenomena interactions between different spins and their environment were neglected. In reality, the excited magnetization does not precess infinitely around the z-axis, but tends to return to its equilibrium value and consequently, align itself parallel to the magnetic field. Two different processes are responsible for the relaxation of the excited magnetization. For the following considerations, the magnetization vector will be divided into a longitudinal component $M_\parallel = M_z$ and a transversal component $\vec{M_\perp} = M_x \cdot \vec{e_x} + M_y \cdot \vec{e_y}$, like visualised in Figure 1A.

- **Spin-Lattice-Interactions** cause relaxation of the longitudinal magnetization $M_\parallel$: After applying an RF pulse, more spins are excited and are aligned anti-parallel to the

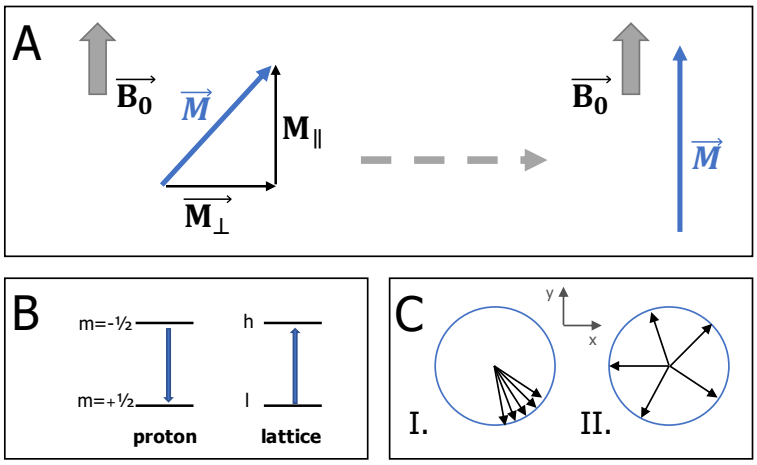


FIGURE 1: A: Visualising the relaxation of the magnetization and illustrating the splitting of the magnetization vector into longitudinal and transversal components, $M_{\parallel}$ and $\vec{M}_{\perp}$ B: Longitudinal relaxation: spins returning to the lower energy level accompanied by an equivalent transition upwards from lattice state l to h. C: Transversal relaxation: different spins fanning out in the transversal plane leading.

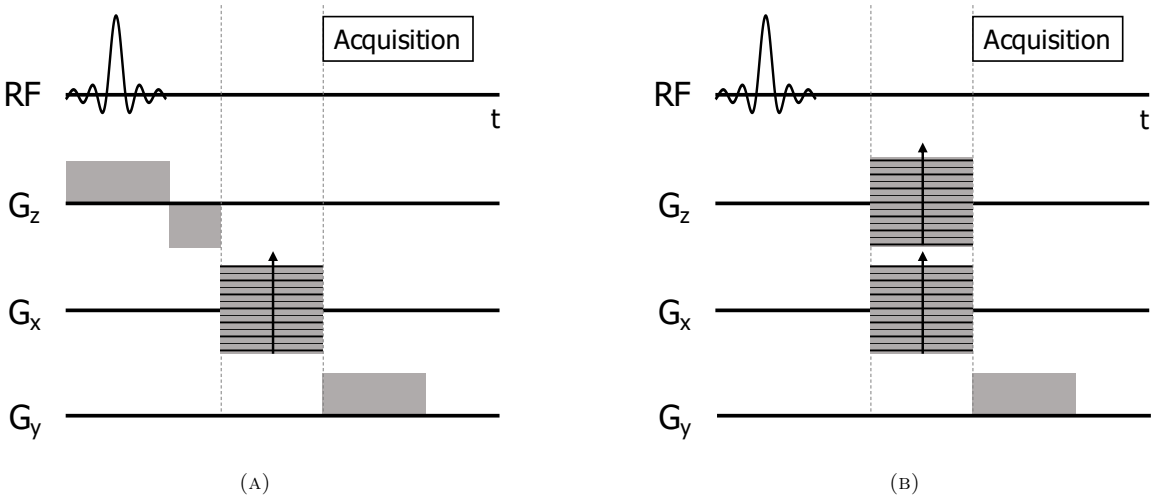
magnetic field, which means that the upper energy level is more occupied than at equilibrium state. However, as the system tries to minimize its energy, spins from the upper energy level return to the lower level by non-radiative transitions, transferring the energy difference to the environment or lattice (cf. Figure 1B). This relaxation of the longitudinal magnetization to its equilibrium value M_0 is described by the time constant T_1, also referred to as longitudinal relaxation time. Brain tissue has a longitudinal relaxation time of approximately 1s, pure water several seconds [18].

- **Spin-Spin-Interactions** lead to a relaxation of the transversal magnetization $\vec{M}_{\perp}$: Within a sample the spins do not only experience the external magnetic field, but also fields caused by surrounding intra- and intermolecular magnetic moments, resulting in time and position dependent local magnetic fields. Thus, different spins precess at slightly different Larmor frequencies, leading to a fanning out in the transversal plane, as visualised in Figure 1C. This dephasing results in cancelling out of the individual transversal magnetic moments and consequently a decay of the transversal magnetization, which is characterized by the time constant or transversal relaxation time, T_2. Brain tissue has transversal relaxation times of around 100 ms, pure water several seconds [18]. Additionally, inhomogeneities of the B_0-field lead to faster dephasing of the magnetic moments, which is summarized in the time constant T_2^*.

2.1.4 Image Formation and Reconstruction

Summarizing the above discussed principles used during an MR experiment, a sample is placed in a homogeneous magnetic field of a permanent magnet or superconducting coil. For excitation of the nuclei, a RF coil transmits another magnetic field, which oscillates at the resonance frequency of the imaged nucleus. After excitation, the magnetization precesses and induces a detectable signal in the receiver coil. The detected signal

$$S(t) \propto \int_V m(\vec{r}, t) \exp(-i\omega_0 t) dV \tag{10}$$



(A) (B)

FIGURE 2: Series of RF pulses and gradients for (A) 2D-encoded and (B) 3D-encoded imaging.

is a damped oscillation, also called Free Induction Decay (FID), with the frequency ω_0 and the time constant T_2^*. $m(\vec{r},t) = M_x + iM_y$ represents the transversal magnetization.

As the aim of Magnetic Resonance Imaging is not to measure the total signal of all nuclei in the sample, but to detect their spatial distribution and consequently, produce a tomogram, the above described set-up needs to be expanded. In order to encode spatial information to the signal $S(t)$, additional magnetic gradient fields are used to modify the signal's frequency and phase [17]. For acquiring 3D data sets a distinction is made between 2D and 3D encoded imaging.

Starting with 2D-encoded imaging, which can also be described as "Multi-Slice 2D Imaging", three different magnetic field gradients, G_x, G_y and G_z with $G_i = \frac{dB_z}{dx_i}$, are used (cf. Figure 2a):

- **Slice Selection Gradient:** In order to only excite one slice of the imaging volume and thus, reduce the 3D to a 2D problem, a gradient G_z is applied during excitation, which will change the resonance frequency of the slices as follows:

$$\omega_0(z) = \gamma(B_0 + zG_z). \tag{11}$$

  A rectangular slice profile with the slice thickness Δz will be achieved using a sinc-shaped RF pulse with the bandwidth $\Delta\omega = \gamma G_z \Delta z$. In the case of a non-180° RF pulse, a gradient with opposite polarity is applied after excitation in order to refocus the dephasing of different subslices. The duration of this refocusing gradient depends on the applied RF pulse.

- **Phase Encoding Gradient:** The phase offset of the magnetization can also be spatially encoded, using a gradient G_x, which is applied before readout for the time T_{ph}. The acquired phase is determined by $\phi(x) = \gamma x G_x T_{ph} = k_x x$, which leads to the magnetization:

$$m(\vec{r},t) = |m(\vec{r},t)|e^{i\gamma(B_0 t_0 + k_x x)}. \tag{12}$$

  As contrast to the frequency encoding technique, the signal is only measured at one time, t_0. However, the excitation and readout need to be repeated with varying phase-encoding (PE) gradients to sample the whole k-space.

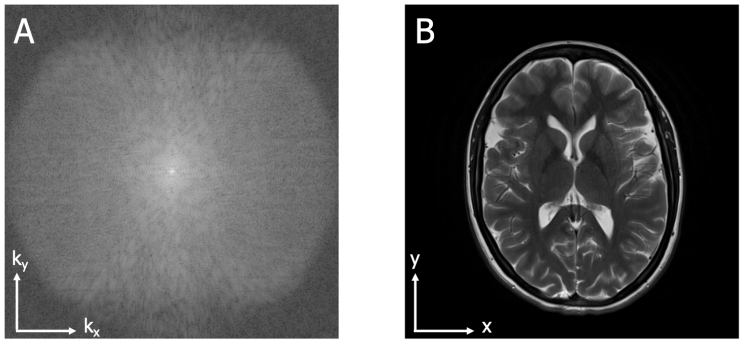


FIGURE 3: (A) Raw data in frequency domain and (B) image in spatial domain.

- **Frequency Encoding Gradient:** By implementing a gradient G_y during readout of the signal, the Larmor frequencies of the single spins are varied according to $\omega_L = \gamma(B_0 + yG_y)$. Consequently, all nuclei induce a signal, oscillating with a frequency which depends on their position along the y-axis:

$$m(\vec{r}, t) = |m(\vec{r}, t)|e^{i\gamma(B_0+yG_y)t}. \tag{13}$$

  By measuring the signal at different time steps within the FID, the frequency space or k-space will be sampled in direction of $k_y = \gamma G_y t$.

Combining equations (13) and (12) with equation (10) determines the signal in the k_x-k_y-space. By discrete 2D inverse Fourier Transformation of the sampled data, the signal intensities in the spatial domain can be calculated for a single slice along the z-direction.

3D-encoded imaging, on the other hand, uses non-selective RF pulses exciting the whole imaging volume (cf. Figure 2b). The slice-selection gradient is replaced by a second phase encoding gradient. Thus, the imaging sequence needs to be repeated $N \times K$ times, with N representing the number of phase encoding steps in x-direction and K the corresponding number of steps in z-direction. The signal intensities in the spatial domain are obtained by discrete 3D Fourier Transformation of the acquired k-space data.

The directions for the respective imaging gradients can be varied depending on the desired image orientation in the clinical setting. Hence, a more general notation will be used for the imaging gradients in the following: G_{SS} for slice selection, G_{PE} for phase encoding and G_{RO} for readout or frequency encoding gradients.

2.2 Motion-related artefacts

As described in the previous section, the MR image is obtained by applying an inverse Fourier Transform to the measured k-space data (cf. Figure 3). Thus, each pixel in the image is a weighted sum of all k-space data points. The center of k-space, relating to low frequencies, contains general image information like contrast and overall shape, whereas the high frequency components in the periphery of k-space represent fine details and information about edges [19]. Subject to the motion pattern and sampling scheme, motion induced changes in one raw data measurement can affect all pixels in the image [20].

For understanding how motion in the spatial domain affects the measurement in frequency domain, the properties of the Fourier Transform can be studied [12]. The Fourier Shift theorem

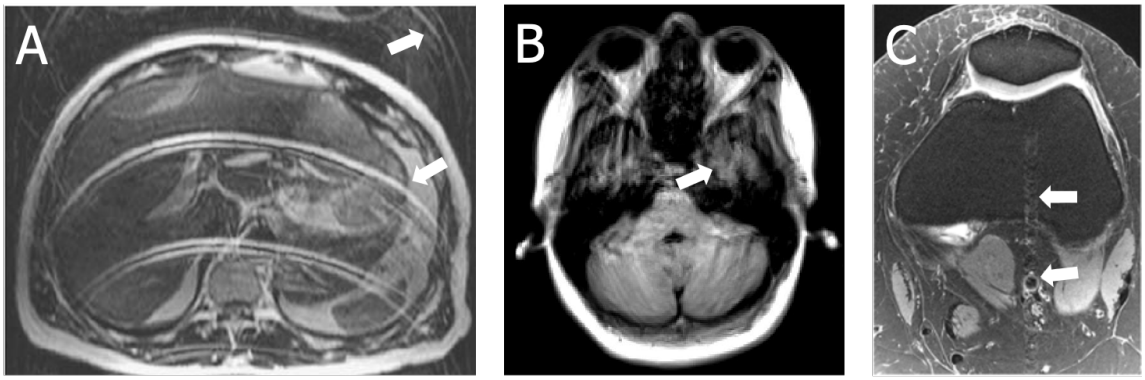


FIGURE 4: Typical motion artefacts: (A) ghosting caused by periodic respiratory motion, (B) blurring caused by random in-plane head motion and (C) flow artefact caused by through plane flow motion. Figures (A) and (C) adapted from [23] and [12].

states that object translation results in a linear phase ramp of the corresponding k-space data in the direction of motion [21]:

$$\mathcal{F}\{I(x - x_0)\} = \exp(-2\pi i x_0 k_x) S(k_x), \tag{14}$$

with the Fourier pair $I(x)$ and $S(k_x)$. Moreover, object rotation introduces an identical rotation in frequency space [21]:

$$\mathcal{F}\{I(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, z)\} = S(k_x\cos\theta + k_y\sin\theta, -k_x\sin\theta + k_y\cos\theta, k_z). \tag{15}$$

Additional effects inducing motion artefacts are phase accumulation, when the object moves during and in the direction of a magnetic field gradient, as well as the *spin history effect*. The latter is caused by through-plane motion, if unsaturated spins with a different excitation history enter the imaging volume. Due to their different magnetization the signal intensity is changed [22].

If the k-space is sampled on a Cartesian grid, motion induced artefacts primarily spread along the phase encoding direction due to the comparably longer acquisition times. Signal samples in frequency encoding direction are acquired within a single echo in 5 to 80 ms, whereas for a complete data set in phase encoding direction, all lines of k-space need to be acquired and hence, the time is in the order of seconds to minutes.

The type of motion influences the induced artefacts [12]. Periodic- and semi-periodic motion such as respiratory motion, flow of cerebrospinal fluid (CSF) or movements caused by swallowing or blinking, result in ghosting. It is defined as the repeated appearance of object structures in the image, as shown exemplary in Figure 4A. In contrast, blurring of details and edges is induced by random motion patterns, like coughing, sneezing or conscious motion of body parts due to discomfort or carelessness (compare Figure 4B). Additionally, signal loss or signal intensity changes occur due to motion induced phase changes which can lead to destructive interference or due to through plane motion. This type of artefact is caused when unsaturated magnetization enters the imaging volume, as for example due to blood flow in a vessel, as shown in Figure 4C.

2.3 Motion Correction

Several methods exist for correcting motion during an MR experiment and thus, avoiding motion artefacts. They can be classified in two categories depending on whether the correction

is applied during or after acquisition. The following two sections give an overview over these prospective and retrospective approaches. For several, but not all of the correction techniques estimates of the relevant motion are needed, which can be determined image-based, RF-based or with external tracking devices. So far, none of the techniques presented below are used in clinical routines.

2.3.1 Prospective Motion Correction

Prospective motion correction is applied during image acquisition in real time. The idea is to preserve a constant relationship between the scanner's field of view (FOV) and the moving object or in other words: "the imaging volume 'follows' the object" [22]. This is realized by constantly adjusting gradient and RF settings based on the motion tracking estimates.

The FOV updates are applied before and sometimes within the individual phase encoding steps. For within echo train correction, fast and high-frequent motion tracking with an external device is superior, since e.g. image-based navigators are not able to measure motion within an individual acquisition block [13].

If severe motion is detected, reacquisition of certain k-space lines can be performed. This is implemented by either reacquiring all lines for which the motion exceeds a certain threshold [24] or by reacquiring a fixed number of k-space lines, ordered by the amount of motion [14, 25].

Prospective motion correction for 3D-encoded sequences has been shown to significantly reduce motion artefacts and improve image quality [13, 14, 24]. It offers several advantages over retrospective methods, like the avoidance of the spin history effect described in the previous section, since updating the FOV prevents through-plane-motion. Such signal intensity changes are extremely challenging for retrospective correction [25]. Additionally, the data is acquired consistently on the predefined sampling scheme and therefore, acquisition is in accordance with the Nyquist theorem. Prospective correction can also be implemented in most of the pulse sequences and hence, offers a large flexibility [22].

Limitations of the prospective approach are the following. Since the correction is performed during acquisition, no uncorrected images are available and images without motion correction cannot be recovered, which makes it more difficult to evaluate the improvement in patient studies. Another challenge of prospective correction is the accuracy of the motion estimates as well as the delay between the physical motion and the time of the update [22]. Furthermore, motion induced changes in magnetic field and gradient inhomogeneities within the object remain uncorrected, since e.g. B_0-shimming is only performed before the scan, which is also a valid limitation for retrospective motion correction.

2.3.2 Retrospective Motion Correction

Different techniques exist for correcting motion retrospectively, that is after image acquisition. They can be divided in machine learning based methods, e.g. using convolutional neural networks like U-Net [26–28] and model-based methods, using motion information or autofocusing techniques [29].

Within this thesis, a model-based approach using motion estimates and raw k-space data is employed, which has been described previously by Slipsager et al. [15]. The rigid-body motion correction is based on the properties of the Fourier Transform (compare equations (14) and (15)). Consequently, translations are accounted for by a phase correction term, which is

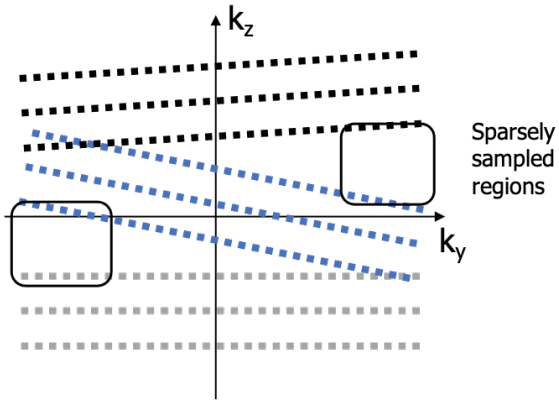


FIGURE 5: k-space correction for rotations around x-axis in spatial domain.

applied to each k-space line separately:

$$S(k_x, k_x, k_z)_{corr} = S(k_x, k_y, k_z)_{raw} \exp\left(-i\Delta\phi_{corr}\right), \tag{16}$$

with $\Delta\phi_{corr} = 2\pi(k_x\Delta x + k_y\Delta y + k_z\Delta z)$. After phase correction, rotation is corrected by rotating the corrupted k-space lines corresponding to the measured object rotation in the spatial domain (cf. Figure 5). As shown exemplary in the figure, after rotational corrections the k-space is not sampled on a Cartesian grid anymore. Therefore, non uniform Fourier Transform techniques need to be used during reconstruction.

This effect represents one limitation of retrospective motion correction, since especially large rotations result in sparse sampling of the frequency space which might introduce artefacts in the spatial domain after reconstruction. Another disadvantage of retrospective motion correction compared to the prospective approach is the fact that artefacts caused by signal intensity changes remain uncorrected. These arise due to the spin history effect or because of parts of the object moving out of the field of view.

TABLE 1: Demographic information, i.e. mean age and sex, for the healthy volunteers participating in the sequence development and the actual study.

	Sequence development	**Actual study**
Sex (F/M)	13 / 5	16 / 6
Age (mean ± std)	26.8 ± 7.6 y	23.5 ± 4.3 y

TABLE 2: Comparison of acquisition times for the different sequences in the original clinical as well as optimised MR protocol for this study.

	Original protocol	**Optimised protocol**	
		(without reacquisition)	(with reacquisition)
T_1 MPRAGE	4 m 26 s	4 m 40 s	5 m 12 s
T_2 FLAIR	2 m 44 s	4 m 12 s	4 m 47 s
T_2 TSE	2 m 27 s	2 m 30 s	3 m 6 s
T_1 STIR	3 m 18 s	3 m 10 s	3 m 51 s
DWI	2 m 7 s	42 s	n.a.
SWI[1] / T_2*	4 m 54 s	2 m 25 s	n.a.
Total time	**19 m 56 s**	**14 m 19 s**	**20 m 3 s**

3 Methods

3.1 Experimental Setup

Within this project, MR and motion data were collected for 22 healthy volunteers between January and March 2021. We acquired the data in a team of two to three researchers, since safety regulations at the hospital require at least two operators during an MR scan.

Participants The participants included within this study were healthy controls recruited for the Cimbi database [30], from which all data is freely accessible. The recruitment was performed by advertisement for different research projects, with approval by the Ethics Committee of Copenhagen and Frederiksberg, Denmark (protocol number H-KF-2006-20). Table 1 summarises demographic information for the 22 participants included in the actual study as well as for the 18 participants recruited for sequence development in the period from September 2020 to April 2021. Written informed consent was obtained from each participant.

MR protocol The protocol consisted of six sequences with different contrast, described in detail in section 3.2. We optimised this protocol based on a clinical cerebral protocol for children. Acquisition times are compared in Table 2. Four scans were acquired per sequence. As illustrated in Figure 6, for the first two scans, the subject was instructed to lie as still as possible. During the third and the fourth scan, we guided the participants to perform a continuous nodding motion for 40 seconds. For one of the sequences (the 3D-encoded MPRAGE) we instructed the volunteers to perform a continuous shaking motion during two additional scans. Prospective motion correction (pMC), was activated for the second of each still and moved scan.

[1] We substituted the susceptibility weighted imaging (SWI) sequence from the original protocol with a shorter T_2*-weighted sequence in the optimised protocol.

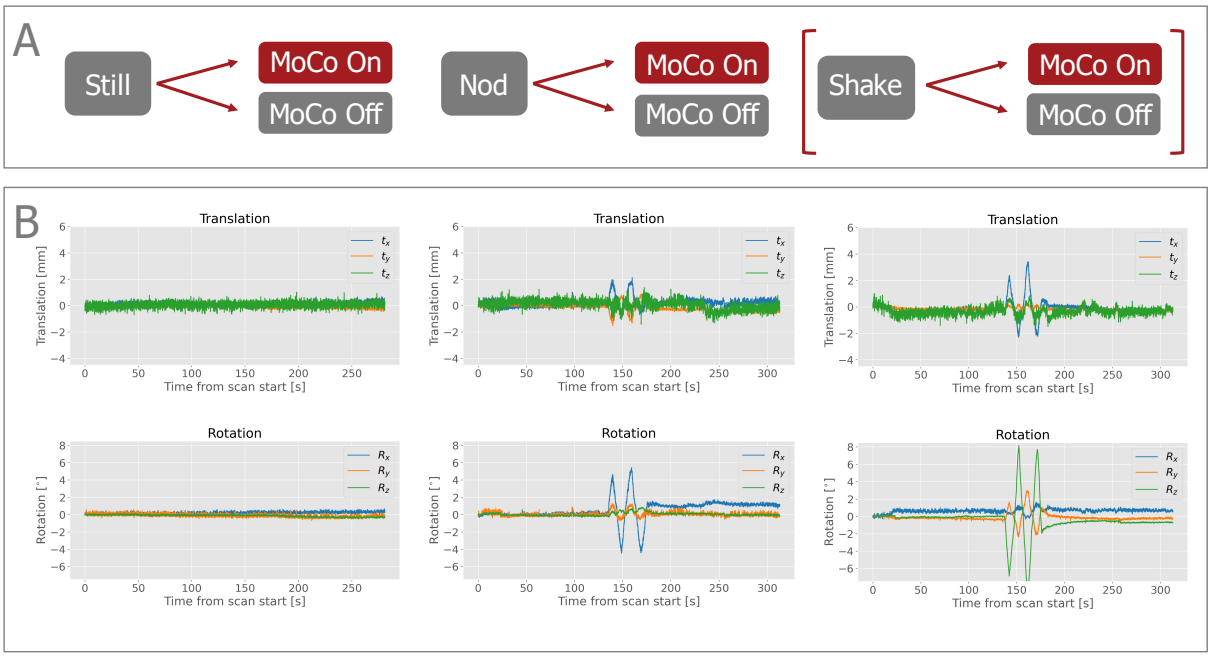


FIGURE 6: (A) Schematic overview of experimental setup; "MoCo On/Off" describing whether prospective motion correction is activated or not. (B) Examples of motion curves for MPRAGE scans with each motion pattern.

The prospective correction algorithm was implemented in the sequences by our collaborators S. R. Frost and A. van der Kouwe (Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts, US). Motion tracking estimates were filtered with a median filter before applying the correction. For the motion acquisitions of four sequences (MPRAGE, FLAIR, TSE and STIR) we selectively reacquired a fixed number of slices at the end of the respective scan, independent of whether pMC was applied or not. For that, the slices with the largest amount of motion, as identified by the tracking data, were chosen. After scanning, we reconstructed the same scan once with and once without the reacquired slices using the program *Twix* on the MR scanner. I will refer to those scans as scans with and without reacquisition in the following.

Motion pattern We chose nodding as main motion pattern based on a previous study [31, 32], where we analysed children's motion patterns in the MR scanner with and without GA. We found significantly more translation along the z-axis, which is the axis pointing into the scanner bore towards the top of the patient's head, compared to the other axes. Furthermore, we were able to show that for children without GA the rotational component around the x-axis was significantly larger than the other components. In the anatomical coordinate system Right-Anterior-Superior (RAS), the x-axis points from right to left. Consequently, the significant rotational component corresponds to nodding motion.

Due to time constraints on the scanning protocol (not considerably more than 90 minutes), we were not able to test the robustness of the motion correction techniques towards a second motion pattern for all sequences. Nevertheless, we chose to perform shaking motion at least for the MPRAGE scan.

The motion was consistently timed to be during the acquisition of the k-space center, in order for the motion to have maximal impact on the acquired image. We guided the volunteers through the movement by telling them the time in 5 s-intervals. For instance for the nodding motion "start, 5, 10, 15, 20 s etc." corresponded to the head being in the center, top, center,

TABLE 3: Definition of ordinal quality scores ranging from 1 (non-diagnostic) to 5 (perfect quality), based on the Likert score.

Score	Description
5	Perfect scan without artefacts.
4	Mild artefacts, but fully diagnostic.
3	Mild to moderate artefact. If an abnormality was noticed, it would be repeated.
2	Substantial artefacts. Large abnormalities can be appreciated, but it has to be repeated.
1	Completely non-diagnostic.

bottom, center position etc. We tried to keep the nodding motion amplitude constant between 8 and 15 mm and the shaking motion amplitude between 12 and 20 mm for all volunteers. This was realized by training the motion before starting the MR scans and monitoring the motion amplitude during the scans. If the motion significantly exceeded the threshold, we repeated the scan.

Retrospective motion correction Retrospective motion correction (rMC), was applied after image acquisition for the two 3D-encoded sequences, as described by Slipsager et al. [15], using the open source software package retroMoCoBox [33]. This was performed by our collaborator J. Slipsager (Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark). I provided him with the raw k-space as well as motion tracking data and he returned the reconstructed nifti images for the two 3D-encoded sequences. However, the retrospective reconstruction algorithm does not include bias field correction, which we then applied afterwards, using the software package *SPM* 12 [34].

Image Quality Scores Additional to the quantitative evaluation of the resulting images, which I describe in sections 3.3 and 3.4, the images were assessed qualitatively by two radiography students based on the ordinal scale presented in Table 3, which is oriented towards the Likert score defined by Kecskemeti et al. [35]. The given scores were in agreement for 53%, 60%, 67% and 63% of all T_1 MPRAGE, T_2 FLAIR, T_2 TSE and T_1 STIR images, respectively. The difference was larger than 1 only in 1.3% of all cases. For a combined quality score I averaged their individual scores.

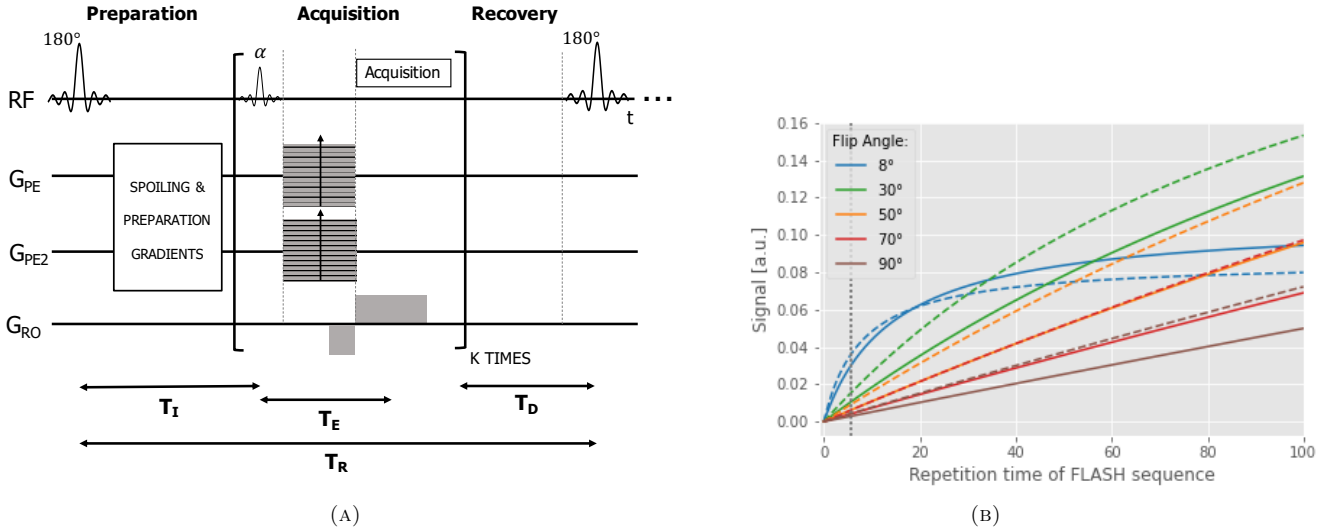
3.2 MR Sequences

The following paragraphs describe the different MR sequences used during data acquisition. All sequences originate from a clinical protocol used for cerebral MRI scans of children. For the T_2 FLAIR, the T_2* and the DWI sequences, the implementation of pMC as well as some imaging parameters needed to be updated during the study. Thus, for these sequences scans from fewer volunteers are available[2].

3.2.1 Magnetization Prepared Rapid Gradient Echo sequence

A Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence is used in clinical protocols to acquire high-resolution 3D isotropic scans with strong T_1-weighted contrast.

[2]More precisely, we performed 19 scans of the T_2* sequence and each 10 scans of both T_2 FLAIR and DWI.



(A) (B)

FIGURE 7: (A) Pulse scheme for the MPRAGE sequence. Please refer to section 2.1.4 for descriptions of G_{PE} and G_{RO}. G_{PE2} represents the second phase encoding gradient. T_R, T_I, T_E and T_D describe repetition, inversion, echo and delay time. (B) Dependency of steady state signal intensities of gray matter (solid line) and white matter (dashed line) on repetition time and flip angle of FLASH sequence. The dotted grey line indicates the repetition time used in this study.

TABLE 4: Sequence parameters describing repetition, inversion and echo times, T_R, T_I, T_E, in ms, as well as RF flip angle α in ° and voxel size in mm. The parameters orientation and phase encoding direction define the directions of the imaging gradients. The parameter scan duration represents the total acquisition time in seconds.

	T_1 MPR	**T_2 FLAIR**	**T_1 STIR**	**T_2 TSE**	**T_2***	**DWI**
T_R [ms]	2000	5000	4100	4010	700	3200
T_I [ms]	900	1800	1600	—	—	—
T_E [ms]	2.32	388	8.5	116	20	79
α [°]	8.0	variable	150.0	150.0	20	90?
Voxel size [mm]	0.9×0.9×0.9	1.0×1.0×1.0	0.9×0.9×5.0	0.4×0.4×5.0	0.5×0.5×5.0	1.7×1.7×4.0
Orientation[3]	sagittal	sagittal	transversal	transversal	transversal	transversal
PE Direction	A → P	A → P	R → L	R → L	R → L	A → P

The 3D-encoded sequence consists of three phases (cf. Figure 7a). In the *magnetization preparation phase* a nonselective pulse with a flip angle of 180° is applied, followed by a delay time T_I for contrast manipulation. Gradients are employed to spoil any transverse magnetization. The *Data Acquisition phase* consists of a short T_R Gradient Echo sequence with flip angle α, more precisely a spoiled Fast Low Angle Shot (FLASH) sequence, using a non-selective excitation pulse for 3D volume acquisition. Fig. 7b shows that the steady-state magnetization of a gradient echo sequence with short repetition times of 4 - 5ms is maximized for small flip angles. At the end of the repetition loop, the *magnetization recovery phase* ensures additional contrast control with a variable "dead time" T_D [36]. The parameters we used for data acquisition are summarized in table 4. In the clinical protocol, this sequence is primarily used for evaluation of brain anatomy and morphology, as well as for evaluating bleedings, migration defects and pathological enhancements post contrast [37, 38].

[3]For 2D-encoded sequences, "Orientation" refers to the direction of the slice selection gradient, for 3D-encoded sequences to the additional phase encoding direction.

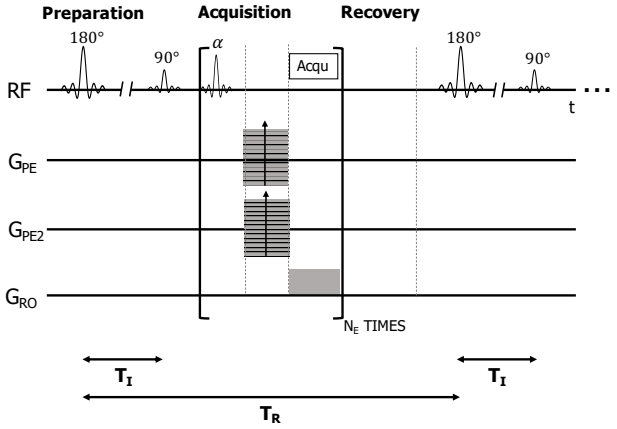


FIGURE 8: Pulse scheme for the 3D FLAIR sequence.

3.2.2 Fluid-Attenuated Inversion Recovery sequence

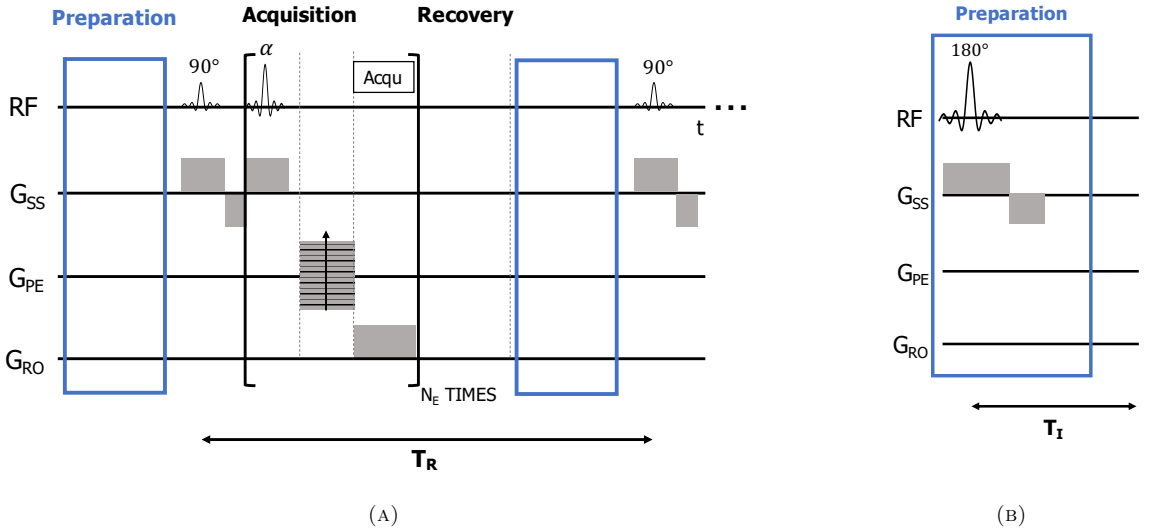
The second isotropic, 3D-encoded sequence used within this work is based on the three-dimensional Fast Spin Echo sequence, on Siemens scanners also called "SPACE" (Sampling Perfection with Application optimised Contracts using different flip angle Evolutions). It consists of a non-selective 90° excitation pulse followed by an echo train of N_E refocusing pulses (cf. Figure 8). These readout pulses have a variable flip angle (typically between 30 and 120°), which ensures high signal at the end of the readout train, as well as reduces tissue heating [39]. The echo time, T_E, for the Fast Spin Echo sequence is defined as the time from the excitation pulse to the acquisition of the center of k-space in the first phase-encoding direction, which is approximately at the middle of the echo train. The base SPACE sequence can be used to acquire images with T_1-weighted, T_2-weighted, proton density or Fluid-Attenuated Inversion Recovery (FLAIR) contrast.

In our protocol, we apply the FLAIR contrast, creating a Siemens based protocol called 'SPACE dark fluid'. As visualised in Figure 8, the magnetisation is prepared by fluid-attenuated inversion recovery, using a non-selective 180° preparation pulse before excitation and acquisition. This magnetisation preparation aims at avoiding artefacts caused by motion of cerebrospinal fluid and at achieving strong T_2-weighted contrast.

The parameters we used for this sequence, are shown in Table 4, as well. Clinically, FLAIR scans are primarily used for detecting white matter signal changes, among others indicating edema, gliosis and tumor inflation [37, 38].

3.2.3 Turbo Spin Echo sequence

Fast- or Turbo Spin Echo (TSE) sequences can be used to acquire 2D-encoded T_2-weighted images. The 2D TSE sequence is a multi-echo version of a spin echo sequence and consists of a 90° slice-selective excitation followed by a series of N_E rephasing pulses with varied phase encoding, as shown in Figure 9a (empty preparation phase). The standard spin echo sequence uses 180° rephasing pulses. However, with multiple echoes and a high flip angle the RF pulses result in a large amount of heating at 3T. Consequently, in order to reduce the specific absorption rate (SAR) the setting *hyper echo* is used, which means that only the first rephasing pulse in each echo train is a 180° pulse, whereas the following pulses have a reduced flip angle α. As described for the T_2 FLAIR, the echo time, T_E, for the Fast Spin Echo sequence is defined as the time from the excitation pulse to the acquisition of the center of k-space in the first



(A) (B)

FIGURE 9: Pulse scheme for 2D-encoded TSE sequences: (A) TSE sequence and (B) magnetization preparation module for FLAIR and STIR sequences.

phase-encoding direction.

The acquisition time is reduced by a factor of $\frac{1}{N_E}$ compared to a basic spin echo sequence [16]. Table 4 again shows an overview of the used parameters. In clinical scans, the TSE sequence is mainly used to evaluate tumors, bleeds, infarcts as well as white matter signal changes [38, 40].

3.2.4 Short-TI Inversion Recovery sequence

A Short-TI Inversion Recovery (STIR) sequence is used to obtain additive T_1- and T_2-weighted contrast with non-selective fat suppression. As Figure 9 shows, it consists of a selective 180° inversion pulse as preparation, followed by a TSE sequence for signal acquisition. Fat suppression is realized by choosing the inversion time T_I such that the longitudinal magnetization of fat spins is nulled at the beginning of acquisition: $T_I = 0.69 \cdot T_{1,fat}$ [41]. The sequence parameters used in our protocol are listed in table 4. Similar to the T_1-weighted MPRAGE, the TIRM sequence is used to assess migration defects, bleedings and posterior pituitary bright spots [42, 43].

3.2.5 T_2*- weighted sequence

T_2*-weighted contrast can be realised with a 2D Gradient Echo (GRE) sequence, similar to the readout of the T_1 MPRAGE. Being a 2D-encoded sequence, it consists of a selective low angle excitation pulse, followed by a combination of dephasing and rephasing gradients generating an echo during readout, as shown in Figure 10. For shortening the repetition time, the remaining transverse magnetisation is destroyed by RF or gradient spoiling. In order to achieve T_2*-weighted contrast with this GRE sequence, the influence of T_1-relaxation needs to be minimised, which is realised by using a small flip angle [17].

In clinical protocols, T_2*- weighted sequences are used for investigating bleeding, hemosiderin in isolation or other pathologies like tumors [37, 44].

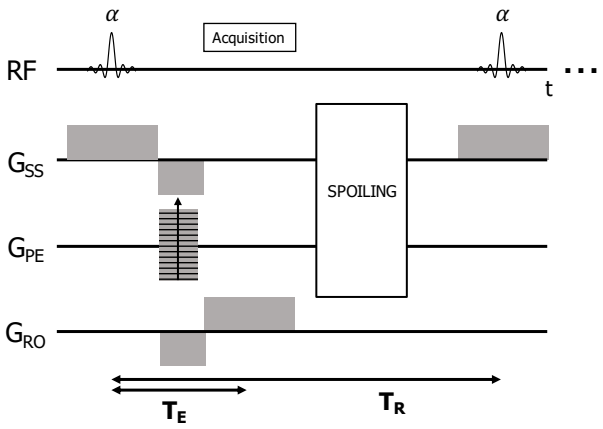
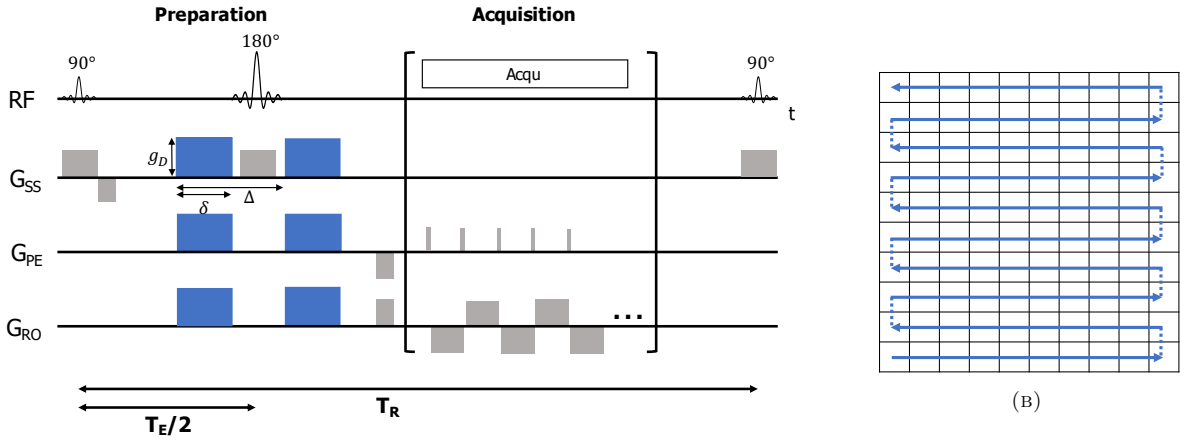


FIGURE 10: Pulse scheme for the 2D T_2*-weighted GRE sequence.



(A)

(B)

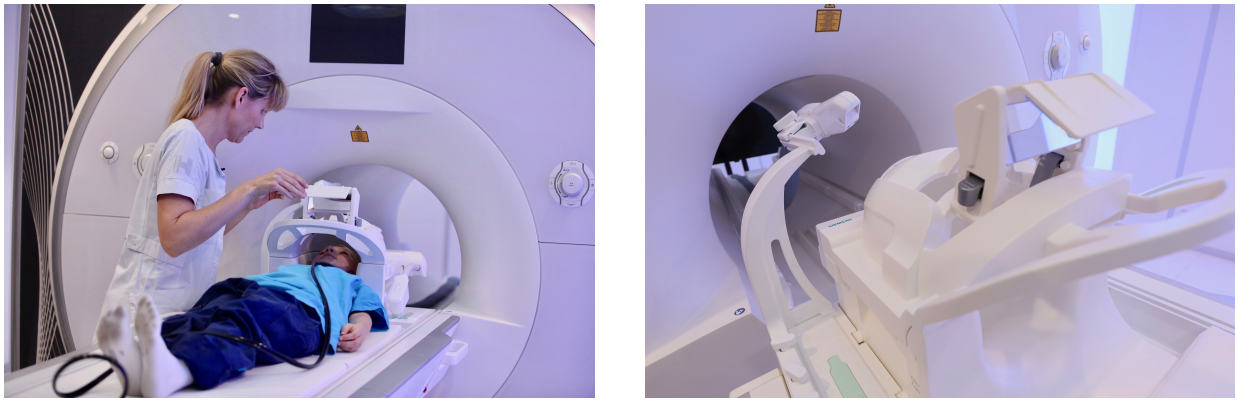
FIGURE 11: (A) Pulse scheme for 2D-encoded diffusion weighted sequence: blue gradients represent the diffusion gradients, which are only active in one direction at a time. (B) Visualisation of EPI k-space sampling pattern.

3.2.6 Diffusion Weighted Imaging

Diffusion Weighted Imaging (DWI) is based on the Brownian motion of water molecules. In tissue, the free diffusion of water molecules is disturbed by e.g. cell membranes and large macromolecules, leading to reduced diffusion described by the tissue dependent apparent diffusion coefficient (ADC) [17].

In order to achieve diffusion weighted (DW) contrast, additional gradients, the so-called diffusion gradients are used in the magnetisation preparation phase of the sequence, as visualised in Figure 11A. In our case, the diffusion gradients are inserted on both sides of the 180° refocusing pulse. The first diffusion gradient introduces additional dephasing of the spins, which is reversed for stationary spins by the second diffusion gradient. For diffusing spins, however, the dephasing caused by the first diffusion gradient cannot be completely compensated by the second gradient, leading to ADC dependent signal attenuation. The diffusion gradients are only activated in one direction at a time.

For signal acquisition we use the echo planar imaging (EPI) technique, which consists of alternating frequency encoding gradients with recurrent phase encoding gradients, leading to



(A) (B)

FIGURE 12: (A) Child lying in the 3T Magnetom Prisma MR scanner. (B) Positioning of Tracoline camera on the scanner bed.

the k-space sampling scheme shown in Figure 11B. Like that, all k-space rows are acquired within one T_R.

In order to quantify the ADC value, at least two diffusion weighted images need to be acquired with different b-values, describing the timing, δ and Δ, and the strength, g_D, of the diffusion gradients [17]:

$$b = (\gamma g_D \delta)^2 (\Delta - \frac{\delta}{3}). \tag{17}$$

In our case $b_0 = 0 \frac{s}{mm^2}$ and $b_{1000} = 1000 \frac{s}{mm^2}$ were used. The ADC map is then given by:

$$\texttt{ADC} = \frac{\log(S(b_0)/S(b_{1000}))}{b_{1000} - b_0}. \tag{18}$$

Clinically, DWI scans are used to evaluate diffusion restriction, T_2 shine through as well as T_2 black out effects and acute cerebral infarction [38, 45].

3.3 Hard- & Software

MR system Magnetic Resonance measurements were performed with a 3 Tesla whole-body MRI system (*Magnetom Prisma, Siemens Healthineers, Erlangen, Germany*) at Copenhagen University Hospital (*Rigshospitalet, Copenhagen, Denmark*), which is shown in Figure 12a. The RF pulses were transmitted and the MR signal was received with a 64-channel head/neck coil, manufactured by Siemens Healthineers *(Erlangen, Germany)*.

Motion tracking We tracked the subject's rigid body head motion with the optical, markerless tracking system Tracoline (*TracInnovations, Ballerup, Denmark*) [46, 47], using the software version *TracSuite* 3.0.245. It projects non-visible infrared light onto the patient's face and reconstructs a 3D point cloud surface, which is continuously tracked at a frequency of approximately 30 Hz. The camera was placed on an arm which can be attached to the scanner table (cf. Figure 12b). The exact position of the camera was fine-tuned before each scan to ensure a good field of view. Frost et al. [13] recommend to ensure that high curvature details of the face, such as the nasal bridge and cheek bone, are within the field of view of the camera to increase the robustness of the motion tracking, like shown in Figure 13a.

(A) (B)

FIGURE 13: (A) Point cloud example visualising the camera field of view of a subject within the head coil. (B) Cross calibration between face selective MR scan (blue) and Tracoline point cloud (white).

Before acquiring any scans of the clinical protocol, a calibration scan needs to be performed, as shown for an exemplar prototype measurement in Figure 13b. This semi-automatic calibration is used to estimate the correct transformation, $A_{tcs2dcs}$, from Tracoline's coordinate system to the device coordinate system of the MR scanner. To transform the motion data into the RAS system, the x- and z-axis need to be flipped, additionally:

$$A_{tcs2ras} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot A_{tcs2dcs}. \tag{19}$$

Data Analysis I performed the analysis of the measured data in Python 3.8.3 (Python Software Foundation). The code is available on `https://github.com/HannahEichhorn/msc-thesis-moco.git`. All MR images were post-processed with the software package *FreeSurfer* v7.1.1 [48], which offers tools for segmentation of brain structures, construction of surface models, as well as intra- and inter-subject registration. In order to extract a brain mask, I applied its function *recon-all*, including 31 steps of skull stripping, white matter segmentation, labelling and surface estimation, to the still, uncorrected 3D-encoded MPRAGE scan. In the following, I will refer to all still, uncorrected scans as "ground truth" scans. I manually examined the extracted brain mask and edited it, in case I observed a wrong segmentation. An example of the brainmask editing is shown in the Appendix (Figure A.2).

Like visualised in Figure 14, I afterwards used the function *bbregister* [49] for registering the remaining MPRAGE scans into the space of the ground truth scan[4], as well as registering the extracted brainmask into the space of the ground truth scans from the other sequences (FLAIR, TSE, STIR, T_2* and DWI). The remaining scans from those sequences were transformed into the domain of their respective ground truth scan with the function *mri_robust_register* [50]. In this fashion, I could apply the same brain mask to all images before calculating the image quality metrics.

In some cases retrospectively reconstructed images showed a ghost-like artefact in the front of the brain, which was not present in images directly reconstructed on the scanner. Figure A.1

[4]The distortion of four MPRAGE shaking scans with retrospective or without MoCo lead to an unsuccessful registration with *bbregister*. These scans were registered with *mri_robust_register*.

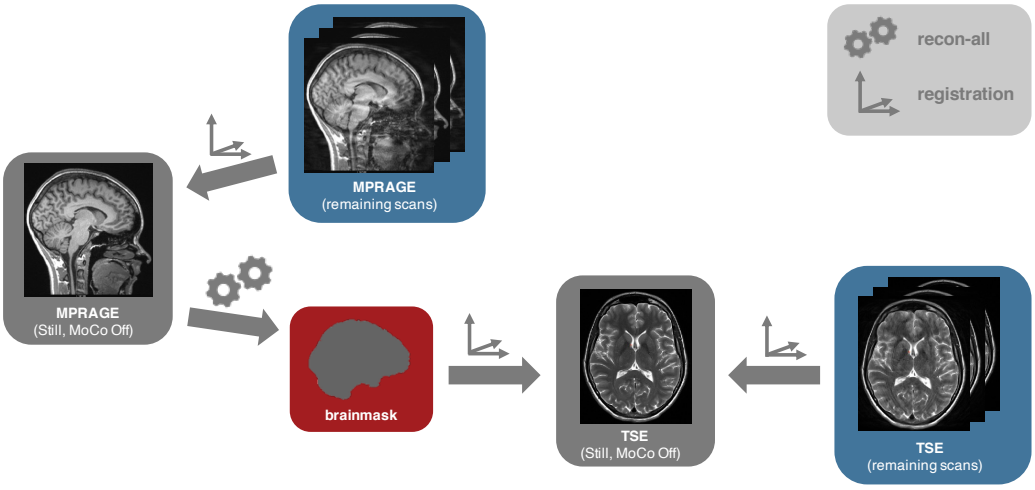


FIGURE 14: Visualisation of registration process with *FreeSurfer* tools *recon-all*, as well as *bbregister* and *mri_robust_register*. The single images represent the still scan without MoCo, the stacked images represent the remaining scans (still with pMC or rMC, nodding and shaking with and without pMC or rMC). The T_2 TSE scans are shown exemplary for all sequences other than the T_1 MPRAGE.

of the appendix can be consulted for an example. Please note that I performed quantitative analyses of image quality, in particular calculation of image quality metrics, exclusively on scans reconstructed off-line in order to not introduce biases in comparative metrics due to differences between the retrospective and the scanner's reconstruction algorithm. For qualitative analyses, in especially observer quality scoring, and for visualisation of image examples in section 4.2, only retrospectively corrected images were used from the retrospective reconstruction. Prospectively corrected as well as uncorrected images were reconstructed by the scanner, in these cases, in order to not introduce negative biases.

3.4 Image Quality Metrics

In order to compare the different motion types as well as motion correction methods, I calculated the following image quality metrics for each scan. All metrics were calculated only for regions within the brain, using the brainmask extracted by *FreeSurfer*. Furthermore, I normalised all scans by subtracting the mean value and dividing by the standard deviation.

Structural similarity index (SSIM) is a reference based metric analysing structural differences between spatially close pixels in a sliding-window approach [26, 51]. It is bounded between -1 and 1 and defined as mean of the values for all image patches $\hat{m}$ and m:

$$\texttt{SSIM}(\hat{m}, m) = \frac{(2\mu_{\hat{m}}\mu_m + c_1)(2\sigma_{\hat{m}m} + c_2)}{(\mu_{\hat{m}}^2 + \mu_m^2 + c_1)(\sigma_{\hat{m}}^2 + \sigma_m^2 + c_2)}, \tag{20}$$

with mean pixel intensities μ and variances σ^2. The covariance between the image patches is defined as $\sigma_{\hat{m}m} = \frac{1}{N-1}\sum_{i=1}^{N}(\hat{m}_i - \mu_{\hat{m}})(m_i - \mu_m)$. $c_1 = (k_1 L)^2$ and $c_2 = (k_2 L)^2$ are variables for stabilization, with $k_1 = 0.01, k_2 = 0.03$ and $L = \texttt{max}(v)$, the maximum value of the reference volume v.

The peak signal-to-noise ratio (PSNR) is another reference based metric, comparing the maximum signal intensity with the noise between image and reference [26]:

$$\texttt{PSNR}(\hat{v}, v) = 10 \cdot \log_{10} \frac{\texttt{max}(v)^2}{\texttt{MSE}(\hat{v}, v)}, \tag{21}$$

TABLE 5: Values of the full-reference metrics SSIM and PSNR for repeated still T_1 MPRAGE and T_2 TSE scans of the same subjects.

Sequence	Volunteer	SSIM	PSNR
T_1 MPRAGE	Prototype 1	0.8567	32.07
	Prototype 2	0.8381	31.07
T_2 TSE	Prototype 2	0.7979	28.07
	Prototype 3	0.7691	27.48

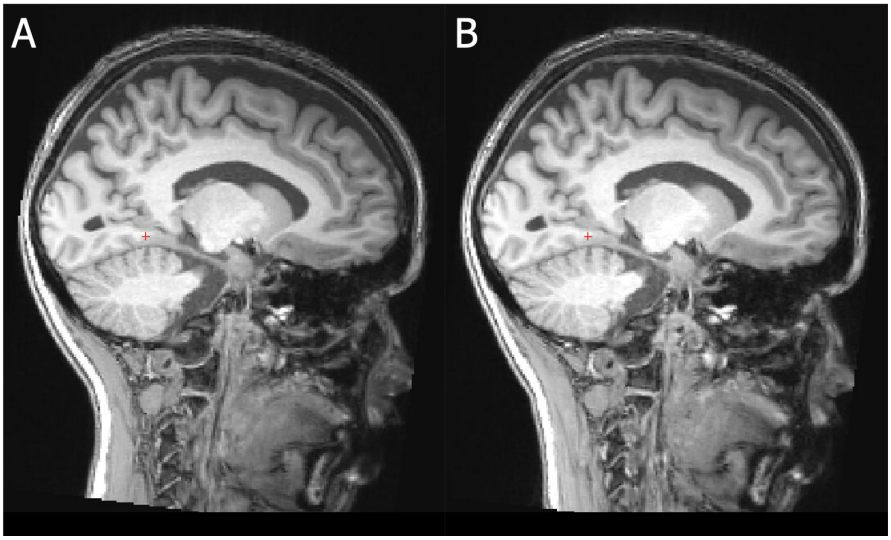


FIGURE 15: Sagittal view of test-retest MPRAGE prototype scans. Image (A) was chosen as reference for calculating SSIM and PSNR of the masked image (B), which are listed as values for Prototype 1 in Table 5.

with the volume of interest $\hat{v}$ and the reference volume v. $\texttt{max}(v)$ represents the largest entry in the reference volume and $\texttt{MSE}$ the mean squared error between the two volumes.

For both SSIM and PSNR, I will use the motionless scan without motion correction as a reference. In order to simplify the interpretation of these full-reference metrics, I calculated exemplary values for two sets of repeated still MPRAGE and TSE prototype scans of the same subject without MoCo in Table 5. Sagittal slices for one set of test-retest MPRAGE scans are shown in Figure 15 to illustrate the small amount of differences, almost invisible for the human eye, which lead to significant deviations of SSIM and PSNR from their maximal values, 1 and infinity, respectively.

The third metric I use to assess image quality within this work is the Tenengrad measure (TGRAD), which analyses image gradients and thus, quantifies sharpness of edges in the image [52]. It is computed (in analogy with the Matlab implementation [53]) by estimating the image gradients G_x, G_y and G_z with a Sobel filter and then averaging all squared gradient magnitudes:

$$\texttt{TGRAD} = \frac{1}{N}\sum_{(i,j)}(G_x(i,j)^2 + G_y(i,j)^2 + G_z(i,j)^2). \tag{22}$$

Since the ADC maps resulting from the DWI sequence represent physical constants, not arbitrary signal intensities, I analysed these maps differently. For each scan except the ground-truth scan, I calculated a difference image from the ground truth scan and then average all absolute values of this difference image. Additionally, following the analysis of Berglund et al. [54], I calculated the mean and standard deviation of the histogram of differences for all scans.

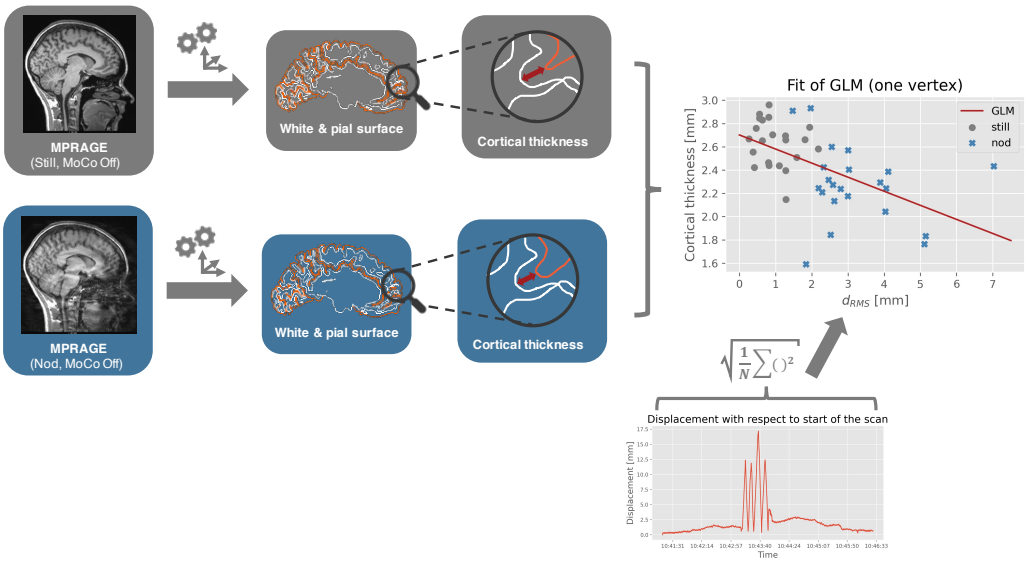


FIGURE 16: Visualisation of the analysis of cortical thickness changes for the 3D MPRAGE scans. The white matter surface is visualised in white, the pial surface in orange. The analysis at the bottom is shown exemplary for nodding motion without motion correction, which can be substituted by all other types of motion or motion correction settings. The plot to the right summarises exemplar data corresponding to a single vertex for all volunteers.

3.5 General Linear Model

Similar to the analyses of Reuter et al. [4] and Frost et al. [13], I investigated the influence of different types of motion and motion correction methods on morphological measures by correlating cortical thickness values with motion.

For generating cortical thickness maps, I processed all MPRAGE scans with *FreeSurfer's* function *recon-all*. Different to the processing of the ground truth for computing image quality metrics, I did not manually edit the brain masks.

In order to compare cortical thickness maps of different subjects, I transformed each subject's map into the space of an average template, *FreeSurfer's fsaverage*. Additionally, I smoothed the resampled maps within the cortex with a 10 mm full-width-half-maximum kernel.

As visualised in Figure 16, cortical thickness is defined as the distance between the white matter surface and pial surface. Blurring of intensities might distort the estimation of those surfaces and thus, falsify cortical thickness values.

I investigated the correlation of cortical thickness values with motion for each vertex by vertex-wise fitting a general linear model to the data of all subjects. As visualised in Figure 16 exemplary for the nodding scan without MoCo, the thickness values of the still scans without MoCo were used as "ground truth" values to compare all remaining sequences with, separately. As quantification of the amount of motion, I calculated the root mean squared (RMS) displacement, d_{RMS}, of all N motion estimates for each scan, using the Euclidean distance with respect to the position at the start of the scan, $\vec{x_0}$:

$$d_{RMS} = \sqrt{\sum_{i}^{N} \frac{(\mid \vec{x_i} - \vec{x_0} \mid)^2}{N}}. \tag{23}$$

The resulting significance maps, derived using the F ratio of explained and unexplained variance, correspond to the p-values for each vertex. A negative p-value indicates cortical thickness

decreasing with d_{RMS} and vice-versa. The significance maps of both hemispheres are thresholded together at FDR= 0.05, (please also see the following section 3.6 for details about multiple comparison correction / thresholding).

3.6 Statistical Analysis

I analysed the differences between image quality measures as well as motion metrics for the different motion correction settings for statistical significance using a Wilcoxon signed rank test. The latter could be applied, since the data was acquired in a paired fashion in this study. A Wilcoxon signed rank test is performed by ranking the absolute differences between the two variables of interest. The test statistic T, which is compared to a table of critical values for determining statistical significance, is then given by the sum of positive or negative differences [55, 56]:

$$T = \mathtt{min}(\Sigma R_+, \Sigma R_-). \tag{24}$$

All tests are performed with a two-sided alternative hypothesis, i.e. testing the null hypothesis of no differences between the two variables of interest ($d = 0$) against positive and negative differences ($\mid d \mid > 0$).

Please note that the Python implementation in the module *scipy.stats* uses the exact distribution for calculating of p-values, if the compared data sets do not include ties. Otherwise, the normal approximation is used [57].

When performing multiple tests, the probability of incorrectly rejecting at least one null hypothesis is considerably larger than the probability of a false rejection for a single hypothesis test. Consequently, since several hypothesis tests were performed within this work, the p-values were corrected using False-Discovery Rate (FDR) [58], at FDR= 0.05.

FDR correction is performed by ordering the m p-values from smallest to largest: $p_{(1)} \leq p_{(2)} \leq ... \leq p_{(m)}$. Then all hypotheses $H_{(i)}$ for $i = 1, 2, ..., k$ are rejected, with k being the largest index for which

$$p_{(i)} \leq \frac{i}{m}\mathtt{FDR}. \tag{25}$$

Please note the different standards for FDR correction: for small sample sizes, such as the comparison of image quality measures, the corrected p-values,

$$p_{(i),cor} = p_{(i)}\frac{m}{i}, \tag{26}$$

are reported, whereas for voxel/vertex based analyses, like the analysis of cortical thickness changes, the original, thresholded p-values are commonly reported. I followed these standards for my analyses.

Estimation of sample size In order to estimate the sample size required to show statistical differences between the motion scans with and without MoCo, we acquired three prototype scans using the MPRAGE sequence and performing a nodding motion. I calculated the above defined metrics PSNR, SSIM and TGRAD for all six scans. The sample size required for a paired t-test at a significance level of $\alpha = 0.05$ and with a power of $1 - \beta = 0.9$ is derived in [59]:

$$n = \frac{21\sigma_w^2}{d^2}, \tag{27}$$

with the mean difference d between the scans with and without motion correction and the within-subject variance $\sigma_w^2 = \sigma^2(1-\rho)$. The latter is calculated with the pooled sample standard deviation[5] σ and the Pearson correlation coefficient ρ. The Pearson correlation coefficient cannot be exactly determined with only three prototype scans, which is why I assumed a modest correlation of $\rho = 0.5$ for our data.

The results of this calculation for the three different metrics are:

$$n_{PSNR} = 18.1 \qquad n_{SSIM} = 7.5 \qquad n_{TGRAD} = 3.5.$$

Consequently, 19 volunteer scans are required for testing significant differences in Peak Signal-to-Noise ratios. Since equation (27) slightly underestimates the required sample size, we aimed for a sample size of 20 volunteers.

[5]For N samples, the sample standard deviation is normalized by $N-1$ instead of N for the population standard deviation.

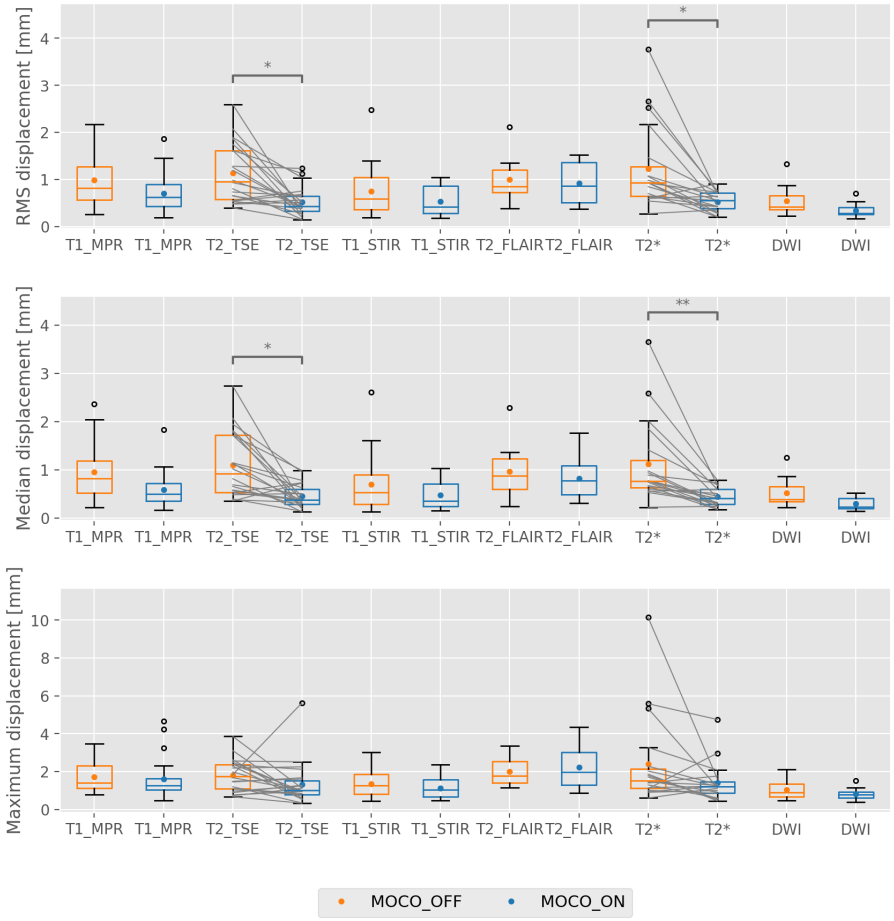


FIGURE 17: Comparison of the RMS, median and maximum displacement between the motionless scans with and without MoCo for the different sequences. Median values of the respective metrics are visualised as box plots, mean values are shown as orange and blue dots. Statistical significance after FDR correction is indicated by */** ($p < 0.05/p < 0.001$). I connected individual paired data points with grey lines for the sequences with statistical significant differences between scans with and without pMC, the T_2 TSE and T_2*.

4 Results

4.1 Motion patterns

In order to compare the image quality of the scans with and without motion correction, it is essential to quantify the amount of motion during the different scans. For that, I calculated three different motion metrics, which are commonly used to quantify motion during MRI [2, 4, 13]: RMS, mean and median displacement. The displacement was calculated as Euclidean distance relative to the start of acquisition of the individual scan. RMS displacement is defined in equation (22).

The values of those metrics for the scans with and without MoCo are compared for the motionless cases in Figure 17 and for the cases with nodding and shaking motion in Figure 18. Please note the different scales of the y-axes in the Figures 17 and 18. In the case of T_2 TSE and T_2* scans without intentional motion, RMS as well as median displacements are significantly larger for the uncorrected compared to the corrected acquisitions. Otherwise, no differences are observed for scans with varying motion correction settings. Median values and standard deviations for these metrics over all sequences are compared for the still and nodding

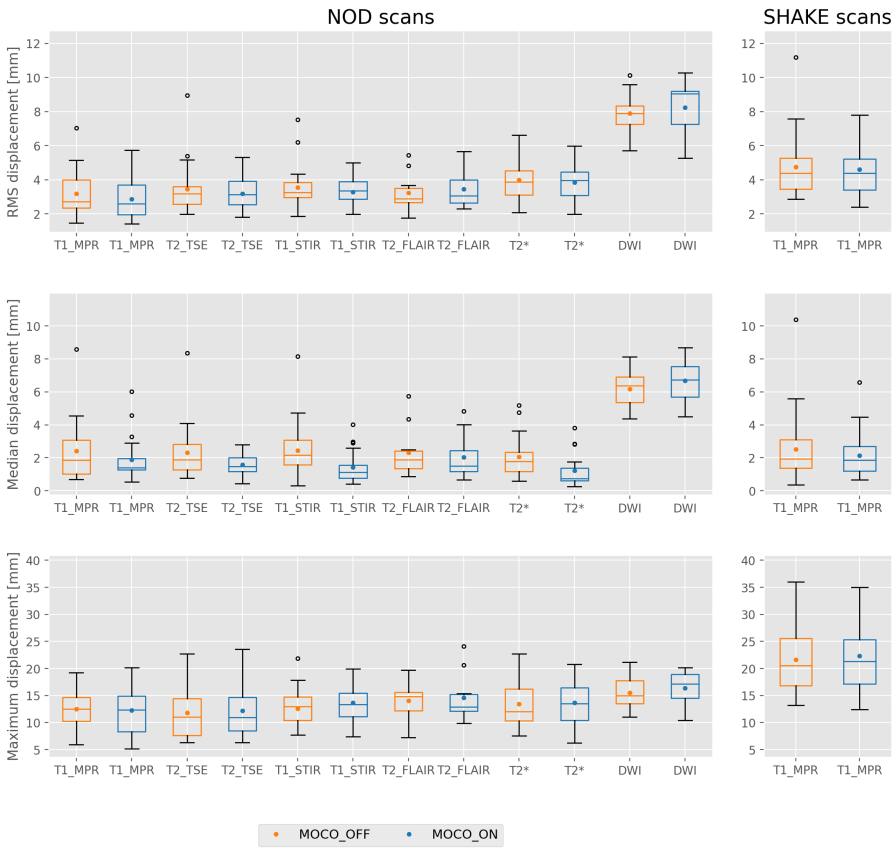
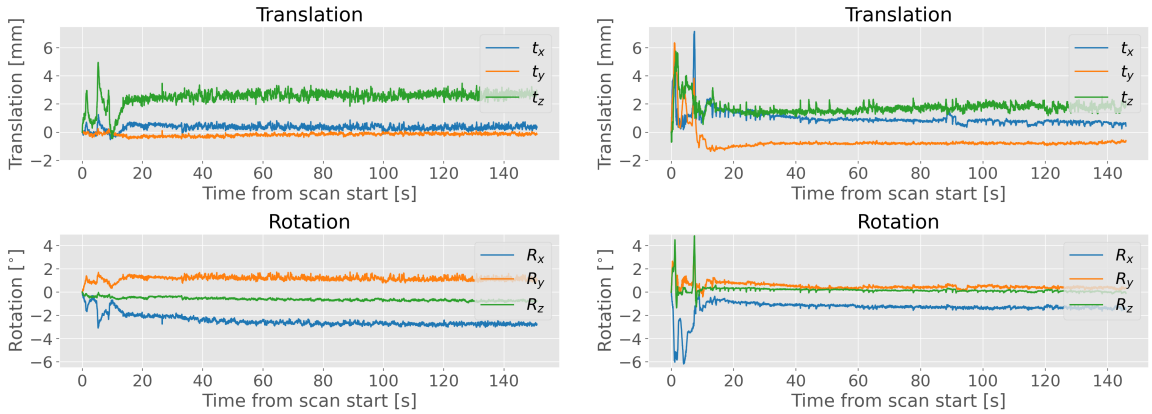


FIGURE 18: Comparison of the RMS, median and maximum displacement between the motion scans with and without MoCo for the different sequences. Median values of the respective metrics are visualised as box plots, mean values are shown as orange and blue dots. No statistical significant differences between uncorrected and corrected scans of the same sequence were observed after FDR correction.

TABLE 6: Summary of motion metrics RMS, median and maximum displacement for still and nodding scans of all sequences.

Motion type	RMS displ. [mm]	Median displ. [mm]	Maximum displ. [mm]
Still (mean ± std)	0.64 ± 0.55	0.54± 0.55	1.3 ± 1.1
Nodding (mean ± std)	3.3 ± 3.0	1.7 ± 2.7	13.1 ± 5.2



(A) Example of T_2 TSE scan.

(B) Example of T_2* scan.

FIGURE 19: Exemplar motion curves for a shift out of the reference position at the start of acquisition of still scans without MoCo.

case in Table 6.

For many still, uncorrected T_2 TSE and T_2* scans I observed some motion in the beginning of the acquisition, corresponding to a shift of the head out of its original position, which I did not observe to the same extent for the still scans with MoCo. Examples for such motion curves are shown in Figure 19.

This shift could possibly influence the metrics based on the displacement with respect to the start of acquisition which are shown in Figure 17. Therefore, I additionally calculated the above described motion metrics with respect to the middle of the acquisition (please consult Figure B.1 of the appendix). In this case, no statistical significant differences could be observed.

4.2 Image quality measures

In the following, I evaluate the influence of prospective and, if applicable, retrospective motion correction on image quality, based on the measures defined in section 3.4, as well as on the observer quality scores described in Table 3. Representative example images are shown for each sequence. In the first subsection I assess the effect of motion correction individually, while investigating the additional effect of reacquisition in the following subsection.

I would like to highlight that the p-values shown in the different sections are always FDR corrected for all comparisons performed for that sequence and not only the comparisons displayed in the respective figure.

4.2.1 Performance of motion correction techniques

Figure 20 shows image examples of the T_1-weighted MPRAGE sequence for nodding scans with and without prospective and retrospective motion correction, next to the still, uncorrected acquisition. Both correction techniques led to a marked improvement. Differences in blurring

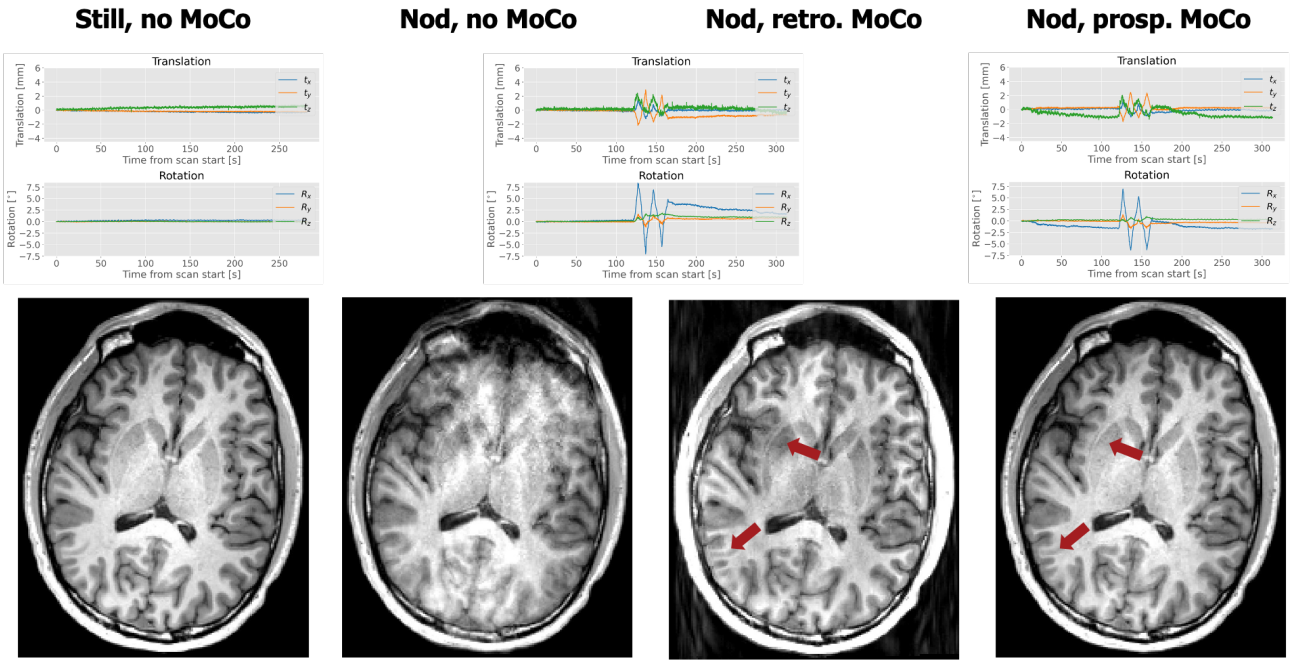


FIGURE 20: Examples of T_1 MPRAGE images without reacquisition, comparing the nodding scans for all motion correction settings with the ground truth scan (left). Contrast was enhanced and blurring reduced for both retro- and prospectively corrected scans. Red arrows indicate regions, where blurring of the scan with pMC was decreased compared to the scan with rMC. The corresponding motion curves are shown above the images. Since rMC was applied to the uncorrected scans, the motion curve in the middle corresponds to both images. All images shown were registered to the still, uncorrected scan.

and contrast between retrospectively and prospectively corrected scans are small and difficult to observe, when they are displayed next to each other. However, the white matter intensities are more uniform for the scan with pMC[6]. The scans with shaking motion showed a similar trend, whereas differences for the different motion correction techniques in the motionless scans were very small (please compare Figures B.2 and B.3 of the appendix).

The metrics and quality scores for both motion types and correction techniques are compared for all subjects in Figure 21. All measures show that prospective motion correction led to significantly higher image quality than retrospective motion correction, which again shows a significant difference compared to uncorrected scans, at least in the scans with shaking motion.

Image examples for rMC and pMC of T_2 FLAIR scans with nodding motion can be found in Figure 22. It is difficult to judge only by visual assessment which motion correction technique performed better in this case, since pMC reduced blurring, but the image shows more ringing artefacts than the retrospectively corrected image. Example images for all motion correction techniques applied to the motionless scans are attached in Figure B.4 of the appendix, which showed slightly more blurring of white and grey matter boundaries for the prospectively corrected scan.

This observation is confirmed by the comparison of image quality measures in Figure 23, which shows significantly smaller values of the reference-free measures for the still scans with pMC. At the same time, marked and significant improvements from nodding scans without

[6]Please be aware that example images with rMC where reconstructed differently than images with and without pMC, as described in section 3.1. Thus, the bias field correction varies for those image type, as well. However, I also noticed the effect of larger uniformity of white matter in prospectively corrected scans, when comparing all retrospective reconstructions of the respective scan types.

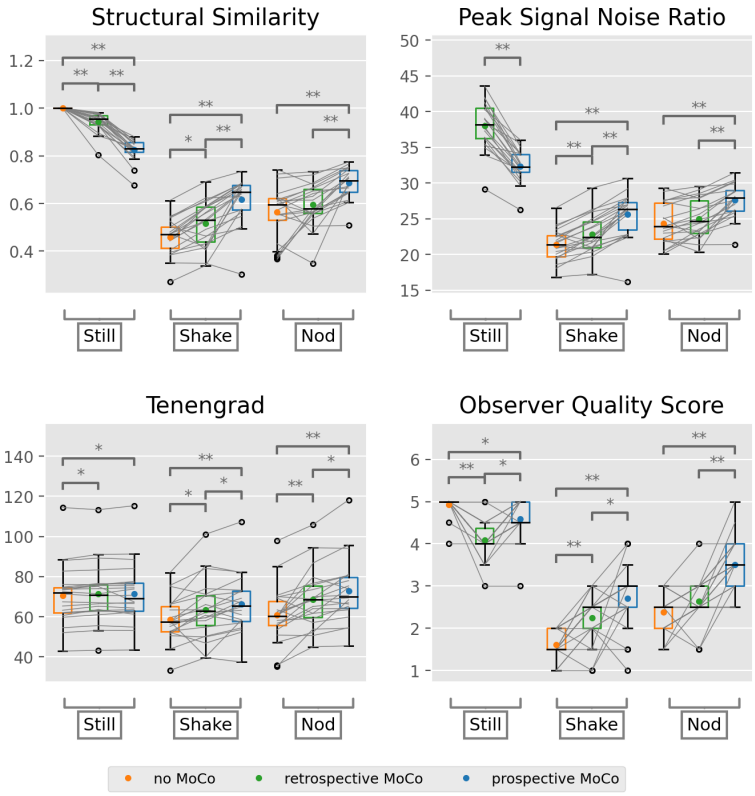


FIGURE 21: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_1 MPRAGE scans with all types of motion patterns and motion correction techniques, without reacquisition. Median values of the respective metrics are visualised as box plots, mean values are shown as colored dots. Statistical significance after FDR correction (of all comparisons) is indicated by */** ($p < 0.05/p < 0.001$). The individual paired data points are connected with grey lines. All image quality measures show a significant increase of image quality from the uncorrected scans, over the ones with rMC to the scans with pMC. Only the retrospectively corrected nodding scans do not vary significantly from the uncorrected scans in terms of SSIM, PSNR and observer quality score. The values for shaking scans are generally lower than for nodding scans.

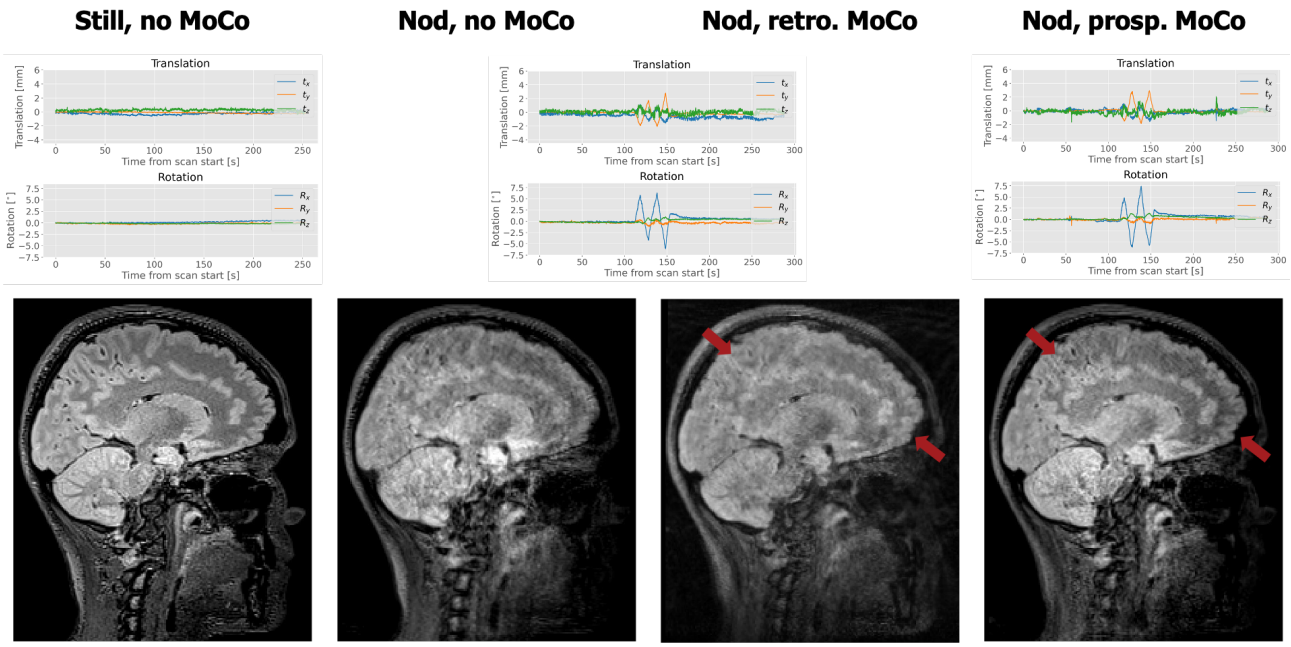


FIGURE 22: Examples of T_2 FLAIR images without reacquisition, comparing the nodding scans for all motion correction settings with the ground truth scan (left). By visual inspection it is not absolutely clear whether pMC or rMC led to higher image quality. As pointed out by red arrows, pMC reduced blurring in frontal parts of the brain, whereas ringing artefacts are less pronounced in the retrospectively corrected scan. The corresponding motion curves are shown above the images. Since rMC was applied to the uncorrected scans, the motion curve in the middle corresponds to both images. All images shown were registered to the still, uncorrected scan.

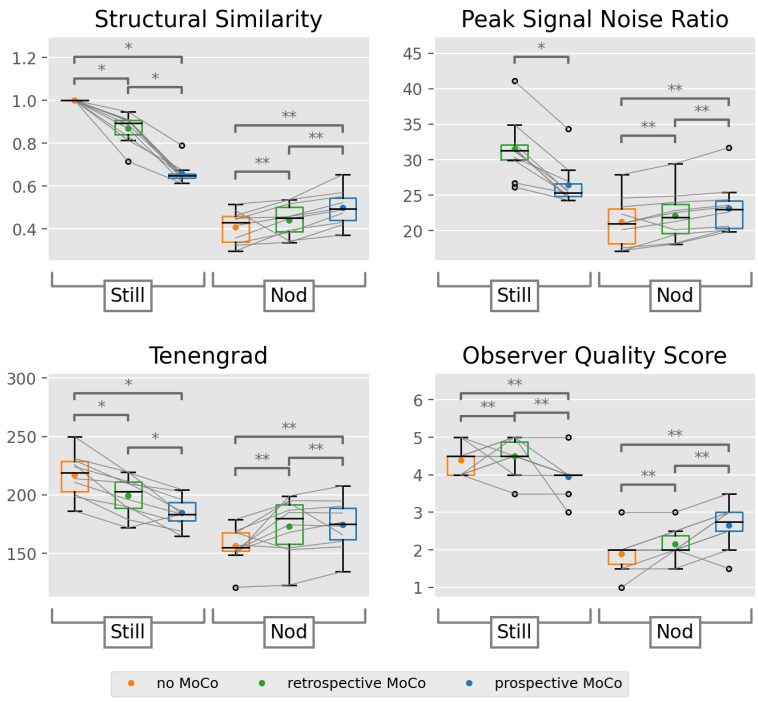


FIGURE 23: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_2 FLAIR scans with all types of motion correction techniques, without reacquisition. All image quality measures show a significant increase of image quality for the nodding scenario from the uncorrected scans, over the ones with rMC to the scans with pMC. Metric values as well as observer scores for still scans with pMC decrease compared to rMC (please consult the caption of Figure 21 for figure details).

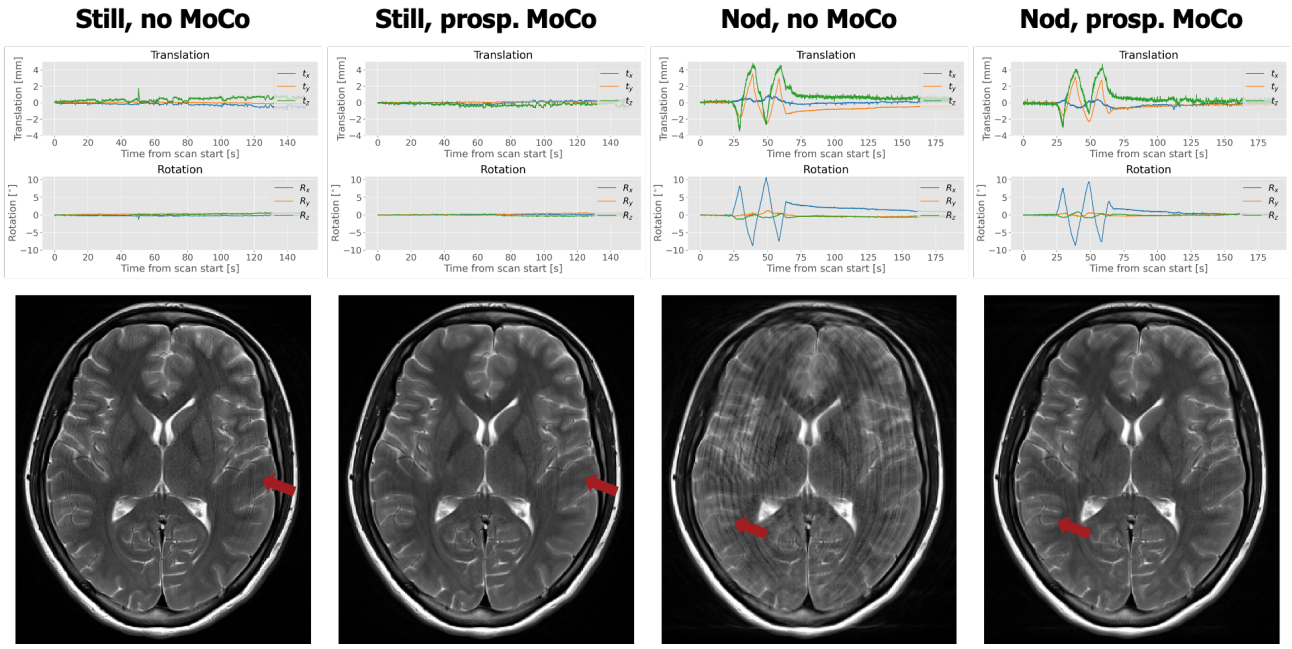


FIGURE 24: Examples of T_2 TSE images without reacquisition, comparing the still and nodding scans with and without pMC. Red arrows indicate regions, where the ringing was reduced in the corrected compared to the uncorrected scan. Especially for the nodding scans, pMC made image structures clearer visible and improved the image quality significantly. The corresponding motion curves are shown above the images. All images shown were registered to the still, uncorrected scan. Please note, for these 2D sequences no retrospective correction was performed.

MoCo, over scans with rMC to scans with pMC are detected.

The examples of T_2 TSE images in Figure 24 illustrate a trend which I observed in the major part of the T_2 TSE scans, namely that prospective motion correction already improved the image quality of the still scan by reducing the amount of ringing. Blurring and ringing is markedly decreased in the corrected image with nodding motion, making the underlying structures clearly recognisable. However, the remaining amount of ringing varied from subject to subject.

The comparison of image quality measures in Figure 25 confirms this observation, showing significant differences for the nodding scans in all metrics.

For the T_1 STIR sequence I could not detect a trend in the results as clearly as for the above discussed sequences. For some subjects prospective motion correction led to reduced ringing and an overall improvement of image quality in the nodding scans, as shown exemplary in Figure 26. Motion artefacts were only present in every second slice. For other subjects the corrected scans did not show reduction in the amount of ringing or blurring. An example of such scans can be found in Figure B.5 of the appendix. In many scans with residual ringing I observed a shift of the head out of the original position after the nodding was performed, which can also be seen in Figure B.5.

Figure 27 confirms this observation to a certain degree. The improvements in image quality due to pMC for SSIM and PSNR are not as pronounced as e.g. for the T_2 TSE scans in Figure 25. In the cases without motion, both Tenengrad measure and observer quality scores show significantly reduced image quality for the corrected compared to the uncorrected scans.

Representative images for the T_2* sequence are shown in Figure 28. Both uncorrected and corrected scans with nodding motion show severe, comparable distortions, which I noticed for

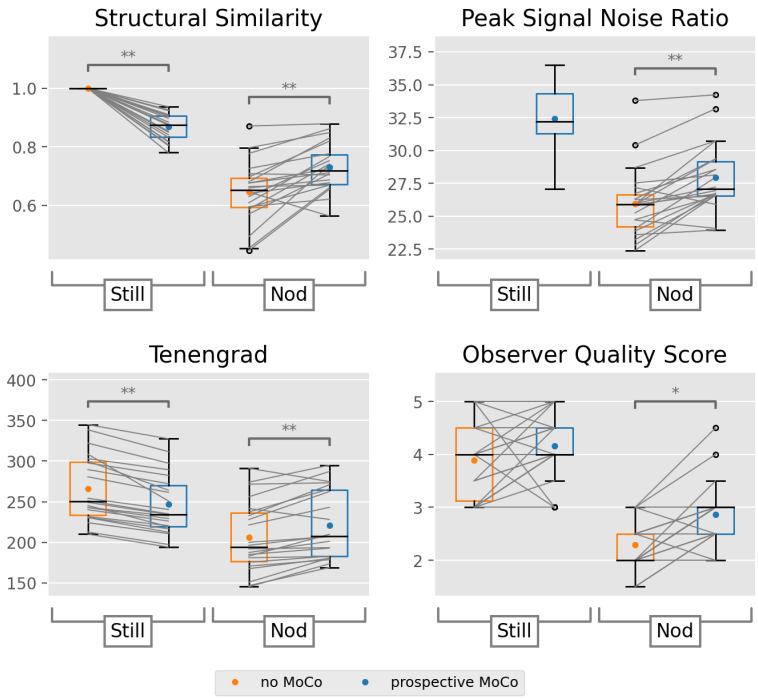


FIGURE 25: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for prospectively corrected and uncorrected T_2 TSE scans without reacquisition. All image quality measures show a significant increase of image quality for the motion scans with pMC compared to the uncorrected scans. Tenengrad values for still scans with pMC are significantly decreased, while observer scores show slightly, even though not significantly, increased quality for still scans with pMC (please consult the caption of Figure 21 for figure details).

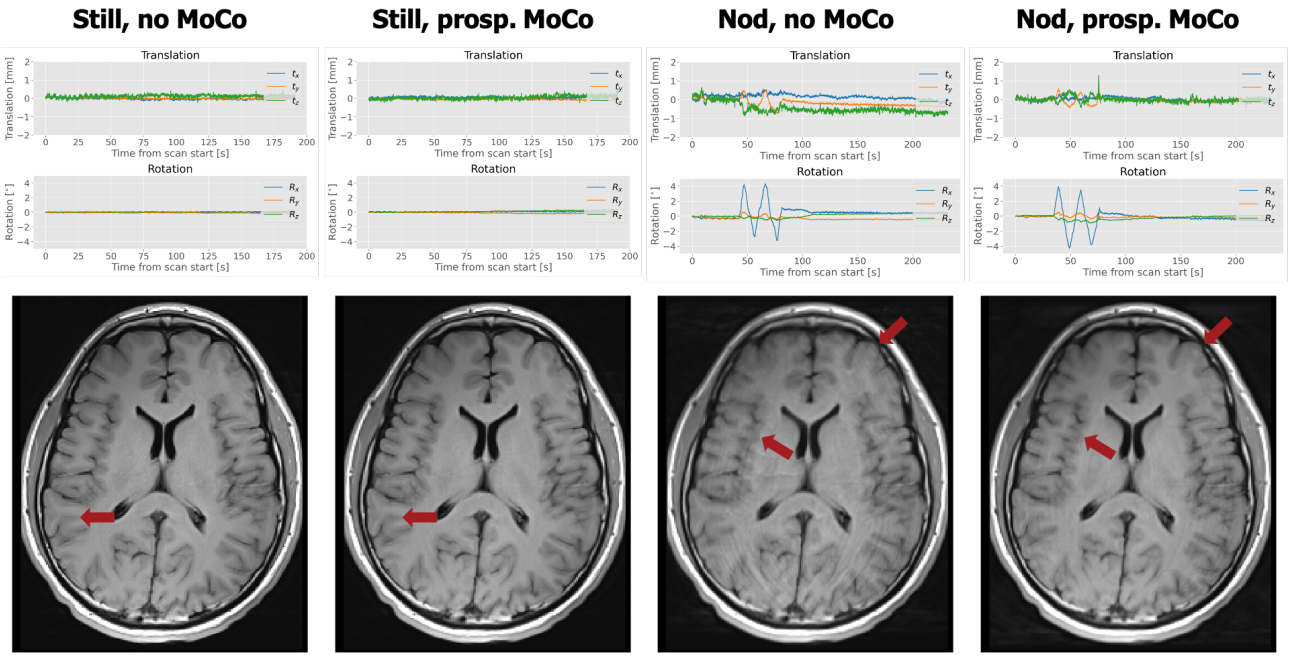


FIGURE 26: Examples of T_1 STIR images without reacquisition, comparing the still and nodding scans with and without pMC. Red arrows in the motionless scans show where pMC led to a small amount of ringing. In the motion scans, red arrows point to areas, where ringing as well as blurring was markedly reduced by pMC. The corresponding motion curves are shown above the images. All images shown were registered to the still, uncorrected scan.

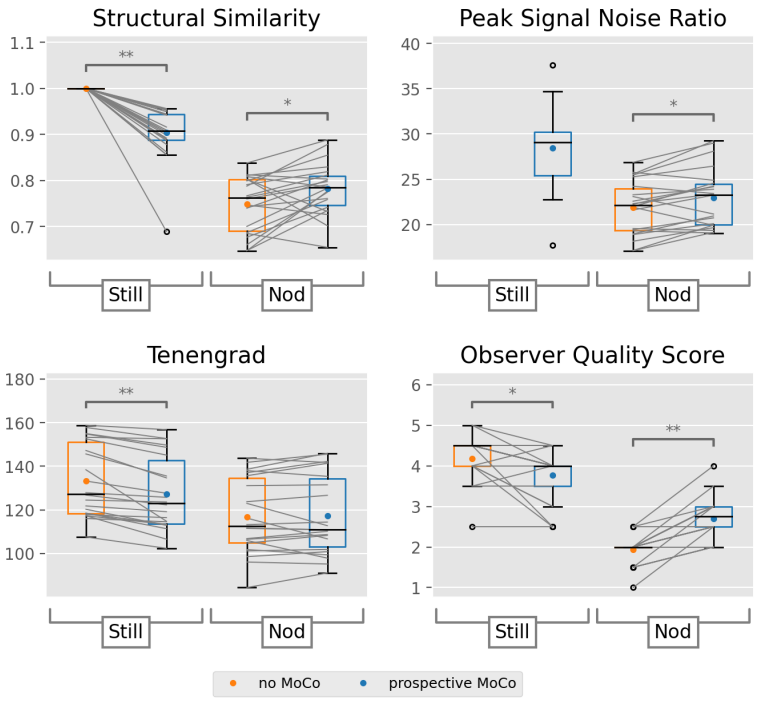


FIGURE 27: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for prospectively corrected and uncorrected T_1 STIR scans without reacquisition. SSIM and PSNR show statistically significant, but comparably small increases in image quality between nodding scans without and with pMC. The values for the Tenengrad measure, however, do not differ significantly for nodding scans, while a small, but significant decrease can be observed for still scans. Observer scores confirm both significant differences in nodding and still scans (please consult the caption of Figure 21 for figure details).

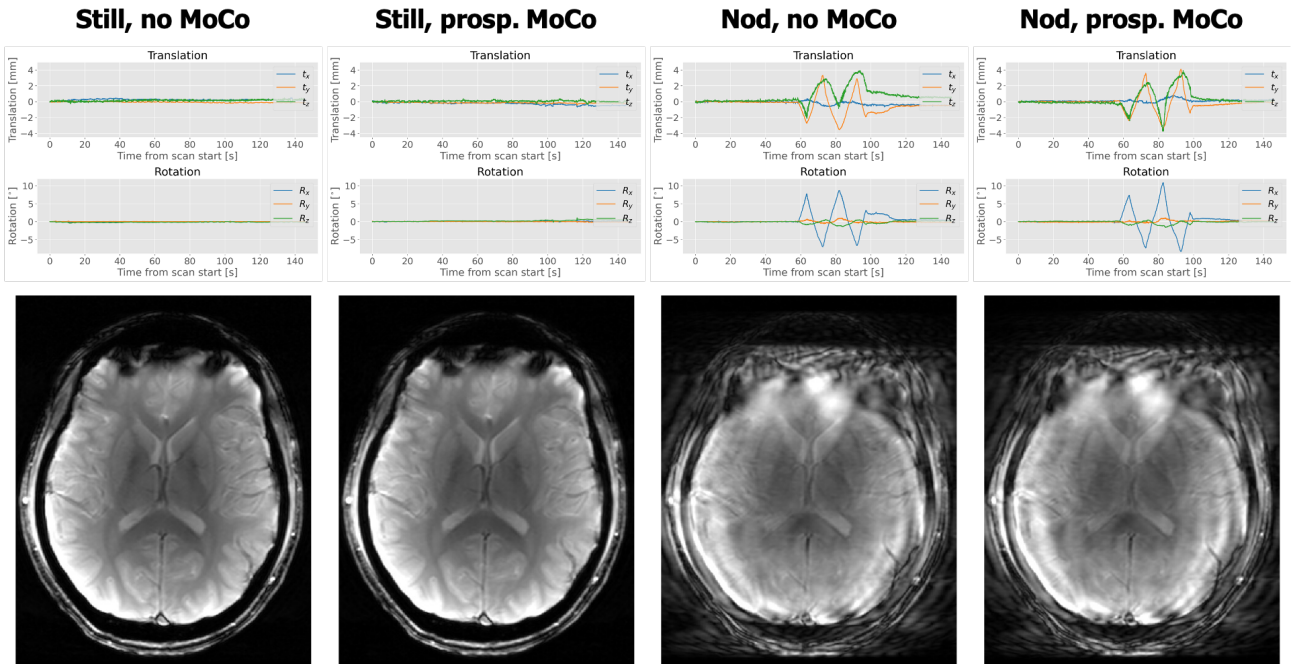


FIGURE 28: Examples T_2* images, comparing the still and nodding scans with and without pMC (no reacquisition). pMC did not improve image quality. No differences are visible for the human eye, comparing the corrected and uncorrected scans. The corresponding motion curves are shown above the images. All images shown were registered to the still, uncorrected scan.

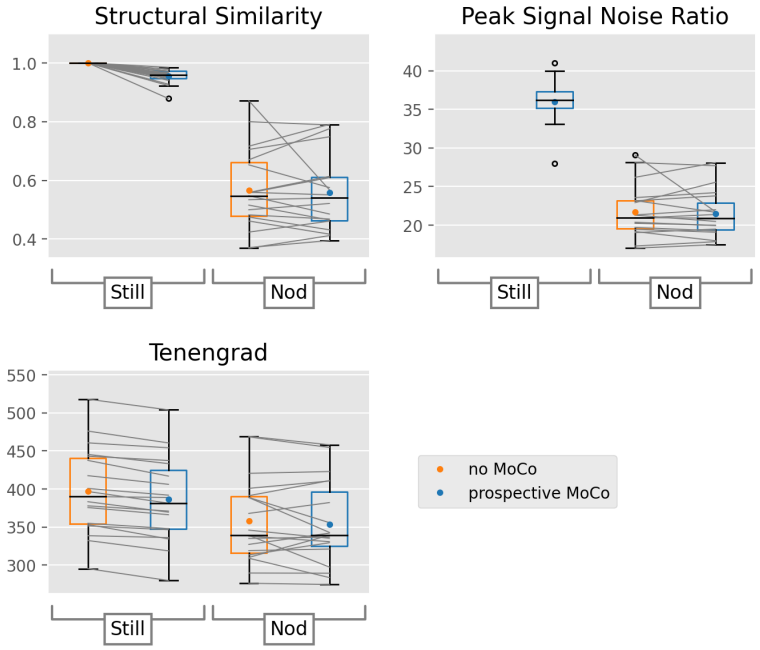


FIGURE 29: Comparison of image quality metrics SSIM, PSNR and Tenengrad, for prospectively corrected and uncorrected T_2* scans (without reacquisition). No significant differences are observed between the scans with and without pMC, the corrected p-values comparing the Tenengrad measure of still scans are borderline, but show a trend towards decreasing image quality due to pMC (please consult the caption of Figure 21 for figure details).

all T2* scans to a similar extend. Consequently, no observer quality scores were given for this sequence, since they are not expected to provide any additional information.

The comparison of image quality metrics in Figure 29 verifies this conclusion. No significant differences are detected between scans with and without prospective motion correction.

Figure 30 shows image examples for the last sequence of our protocol, the DWI sequence. Even though the images were in general not as strongly distorted, I detected small improvements in both TRACE weighted images, as well as ADC maps due to prospective motion correction in the nodding scenario. I observed this trend in the majority of DWI scans.

Quantitative analyses of all TRACE weighted images are shown in Figure 31, which underlines that prospective motion correction improved image quality. Analogous analysis of TRACE weighted images with $b = 0\frac{s}{mm^2}$, showing similar trends, are available in the appendix (Figure B.6). Since ADC maps represent physical constants and not arbitrary signal intensities, I did not calculate the above discussed image quality metrics. Instead, I analysed difference images with respect to the still, uncorrected scan (please compare Figure 32). Significant differences between nodding scans with and without pMC can be observed for mean absolute differences of the ADC maps, as well as standard deviations of the histograms corresponding to each difference map. The means of those histograms did not differ significantly.

4.2.2 Impact of reacquisition

In the following, I evaluate the impact of selectively reacquiring a fixed number of slices with the highest level of motion at the end of the respective scan. The latter was performed for the motion acquisitions of the two 3D-encoded sequences, T_1 MPRAGE and T_2 FLAIR, as well as for two of the 2D-encoded sequences, the T_2 TSE and the T_1 STIR.

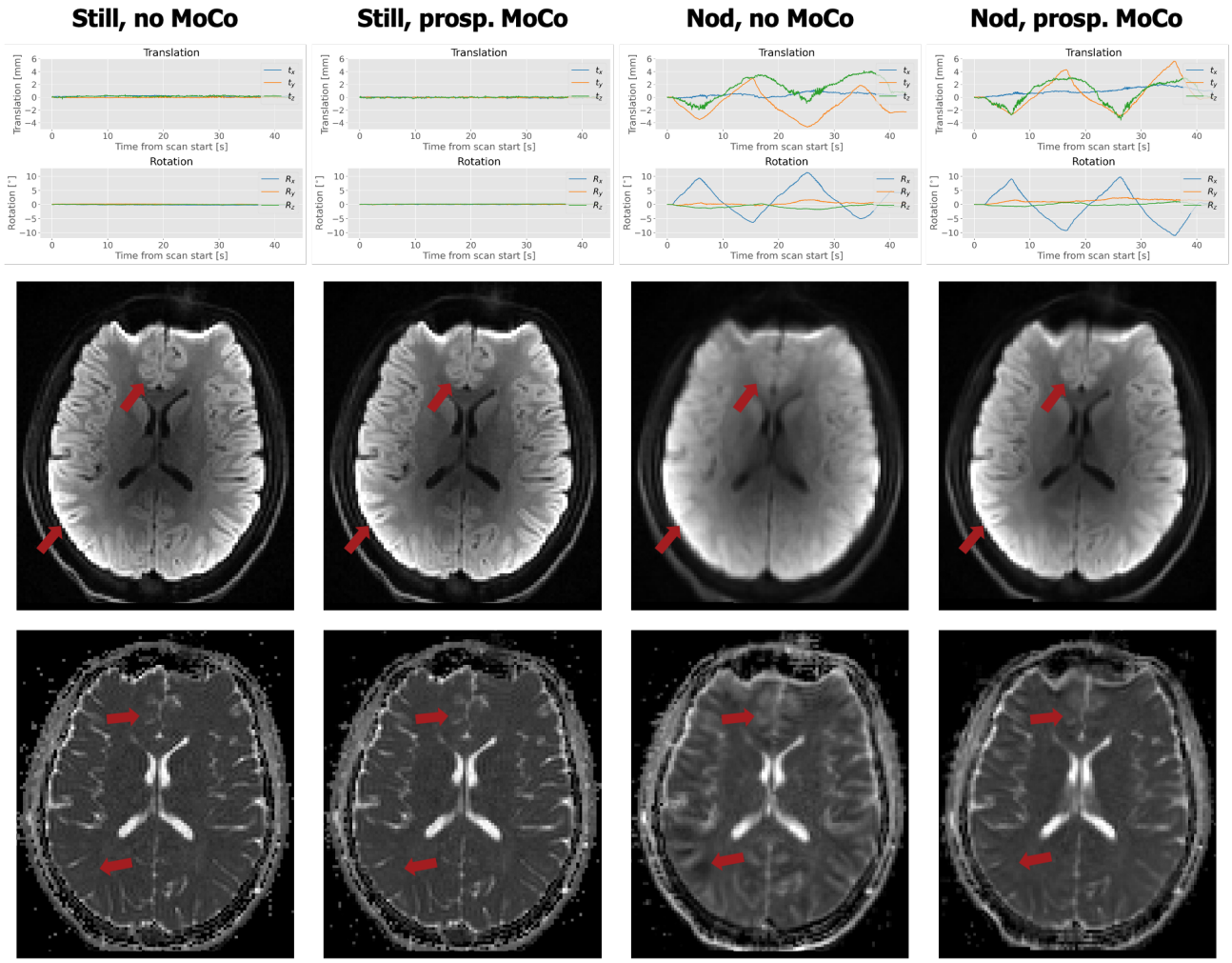


FIGURE 30: Examples of DWI images, comparing the still and nodding scans with and without pMC (no reacquisition). The motion curves for each scan are shown in the first row, TRACE weighted scans with $b = 1000\frac{s}{mm^2}$ are shown in the second and ADC maps in the third. Red arrows for TRACE weighted scans show where blurring was decreased and contrast intensified in the nodding scan with pMC. Similarly, for ADC maps red arrows show reduction of artefacts, more precisely areas, where ADC values of the corrected nodding scan are closer to those of the still scans than for the uncorrected scan. All images shown were registered to the still, uncorrected scan.

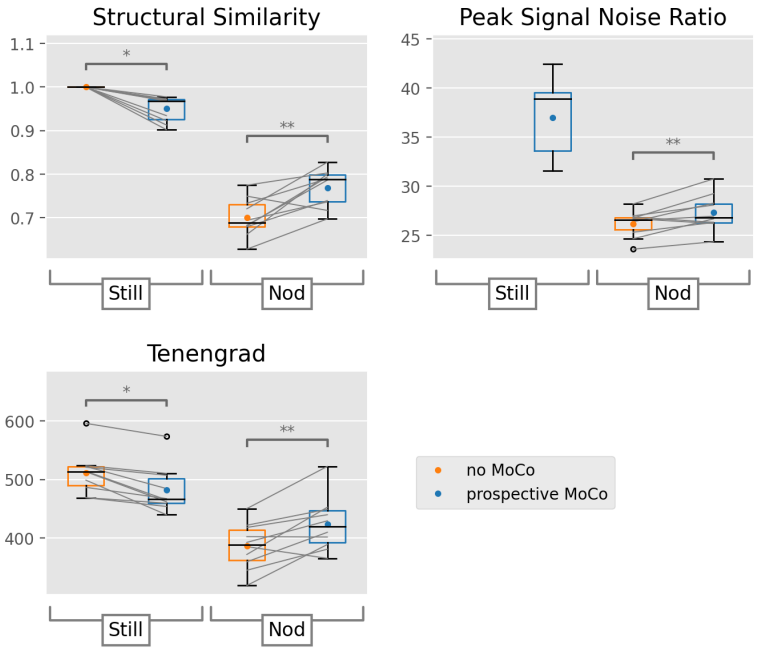


FIGURE 31: Comparison of image quality metrics SSIM, PSNR and Tenengrad for prospectively corrected and uncorrected DWI TRACE weighted scans with $b = 1000 \frac{s}{mm^2}$ (without reacquisition). All image quality measures show a significant increase of image quality for the motion scans with pMC compared to the uncorrected scans (please consult the caption of Figure 21 for figure details).

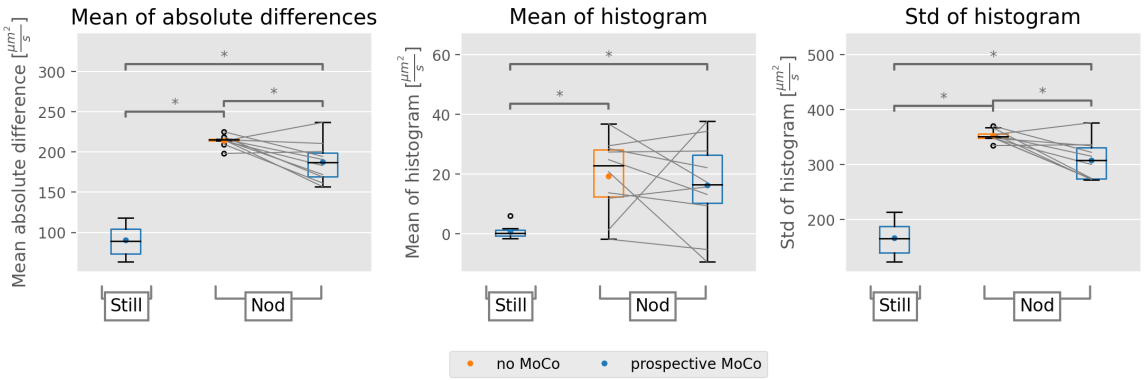


FIGURE 32: Comparison of ADC difference images computed by subtracting the prospectively corrected and uncorrected maps from the still, uncorrected ADC map. From left to right the mean absolute differences, as well as the means and the standard deviations of the corresponding histograms are shown. The literature ADC value for white matter is 700 $\frac{\mu m^2}{s}$ [17]. The nodding scans with pMC show slightly, but significantly reduced mean absolute differences as well as reduced standard deviations, corresponding to higher image quality. No differences are observed in the means of the corresponding histogram (please consult the caption of Figure 21 for figure details).

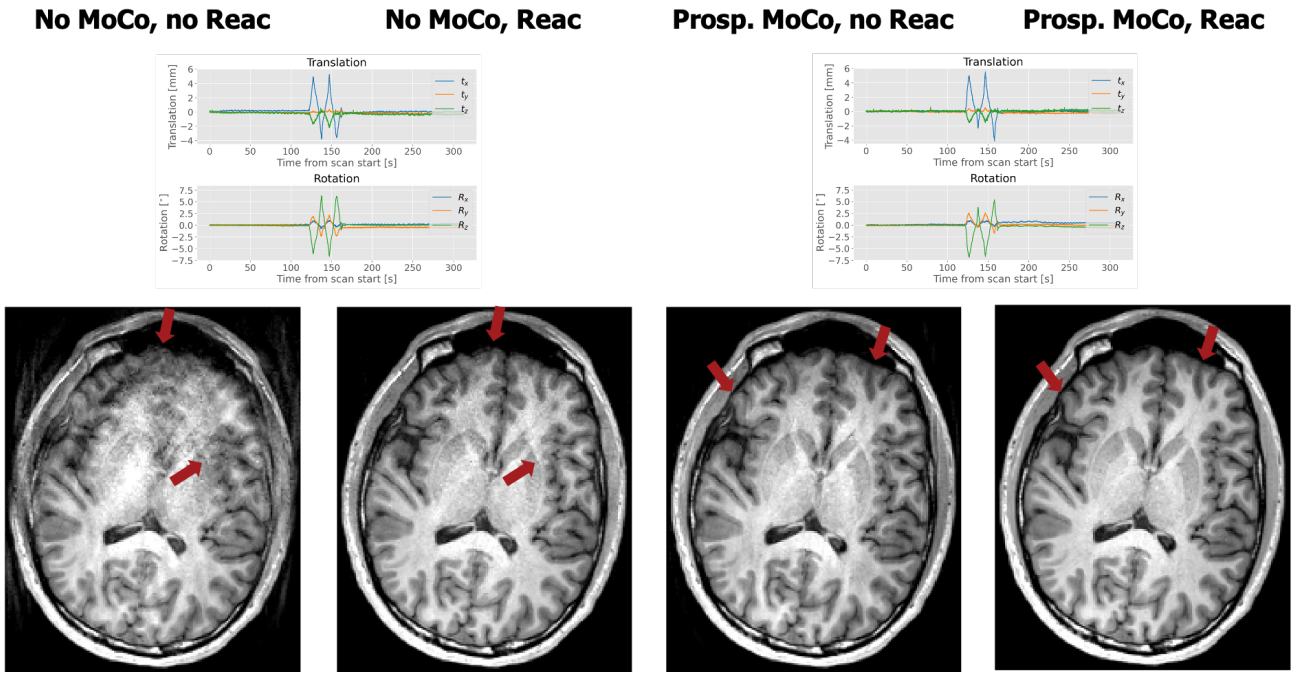


FIGURE 33: Examples of T_1 MPRAGE images with shaking motion, comparing the reconstructed scans without and with reacquired slices for acquisitions without and with pMC. Red arrows indicate regions, where reacquisition decreased blurring of imaged structures. Differences in the scans without MoCo are easer recognisable. Reacquisition led in both cases to more uniform white matter signal intensities. The corresponding motion curves are shown above the two respective images. All images shown were registered to the still, uncorrected scan.

Reacquisition supplementary to motion correction led to additionally increased image quality for the T_1 MPRAGE sequence, which is shown for example images with shaking motion in Figure 33. In Figure 34 I compare image quality metrics and scores for acquisitions with and without prospective motion correction, as well as with and without reacquisition. Please devote particular attention to the large difference between the scans without pMC and without reacquisition, which represent the clinical standard, and the scans with pMC and reacquisition, which represent the best-case scenario. The same analysis for retrospectively corrected scans with and without reacquisition can be found in Figure B.7 of the appendix.

Reacquiring motion affected slices had a large influence on image quality of the T_2 FLAIR sequence, as visualised for image examples in Figure 35. Whereas the prospectively corrected scan without reacquisition is still strongly affected by ringing and blurring artefacts, these can be notably reduced by reacquisition, which I noticed for many subjects. This observation is also reflected in the comparison of image quality metrics and observer scores in Figure 36. Scans with reacquisition outperform the corresponding scans without reacquisition for all analysed measures. The combination of pMC and reacquisition significantly improves the clinical standard, analogously to the results from the T_1 MPRAGE sequence.

Furthermore, I analysed the impact of reacquisition on retrospectively corrected T_2 FLAIR scans in Figure B.8 of the appendix, which shows similar results as Figure 36.

For the T_2 TSE sequence, reacquisition slightly reduced both ringing and blurring artefacts, but the effect is small compared to the improvement by pMC, as visualised for exemplar images in Figure 37. This observation is reflected in the image quality metrics I calculated on scans with and without reacquisition, shown in Figure 38. However, the combination of prospective motion correction and reacquisition led to a clear increase in image quality compared to the

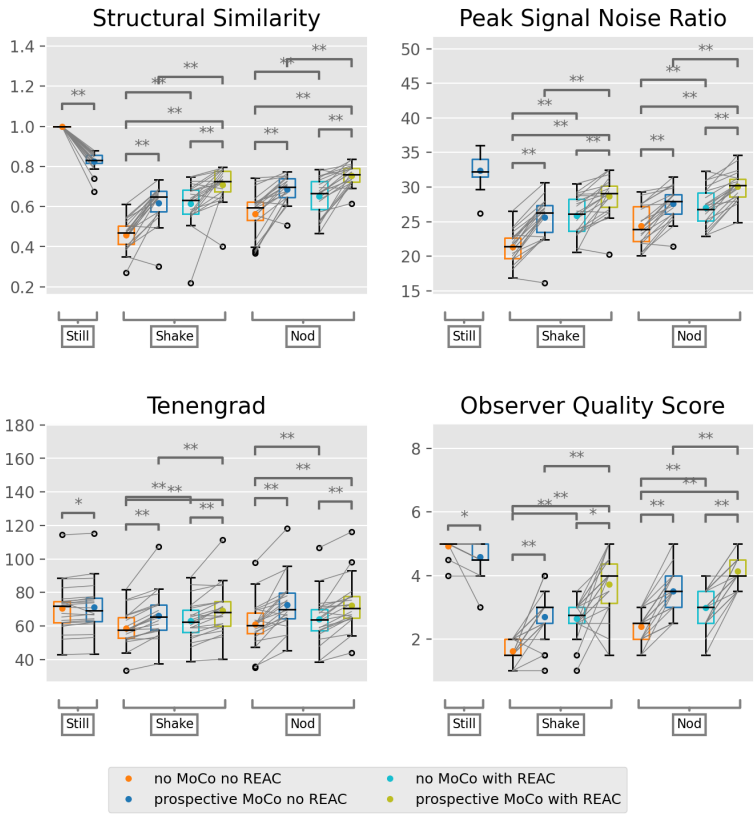


FIGURE 34: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_1 MPRAGE scans with and without pMC as well as reacquisition, for all types of motion patterns. Reacquisition leads to an additional increase of image quality, with particular emphasis on the scans with MoCo and reacquisition (green) which clearly outperform the clinical standard (orange). Median values of the respective metrics are visualised as box plots, mean values are shown as colored dots. Statistical significance after FDR correction (of all comparisons) is indicated by */** ($p < 0.05/p < 0.001$). The individual paired data points are connected with grey lines.

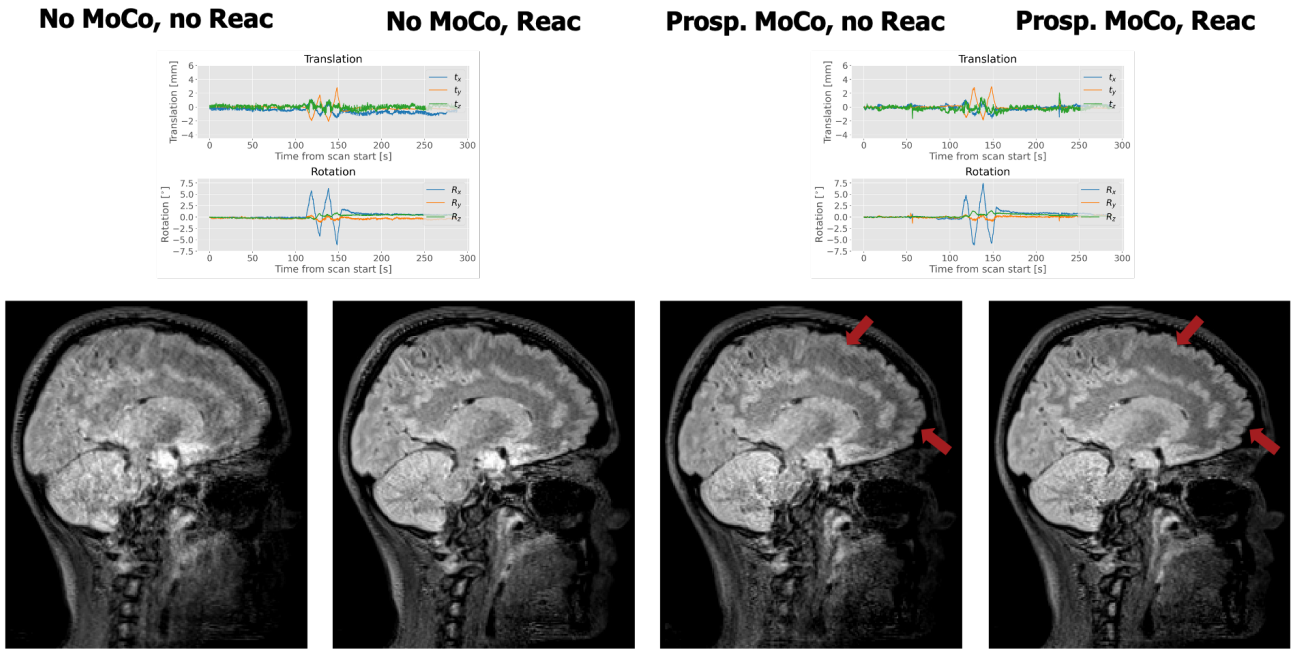


FIGURE 35: T_2 FLAIR image examples with nodding motion, comparing the reconstructed scans without and with reacquired slices for acquisitions without and with pMC. The improvement due to reacquisition, i.e. reduction of blurring and contrast enhancement, is clearly visible for the scans without MoCo. For those with pMC I marked areas with red arrows, where blurring and / or ringing wass decreased in the scan with reacquisition. The corresponding motion curves are shown above the two respective images. All images shown were registered to the still, uncorrected scan.

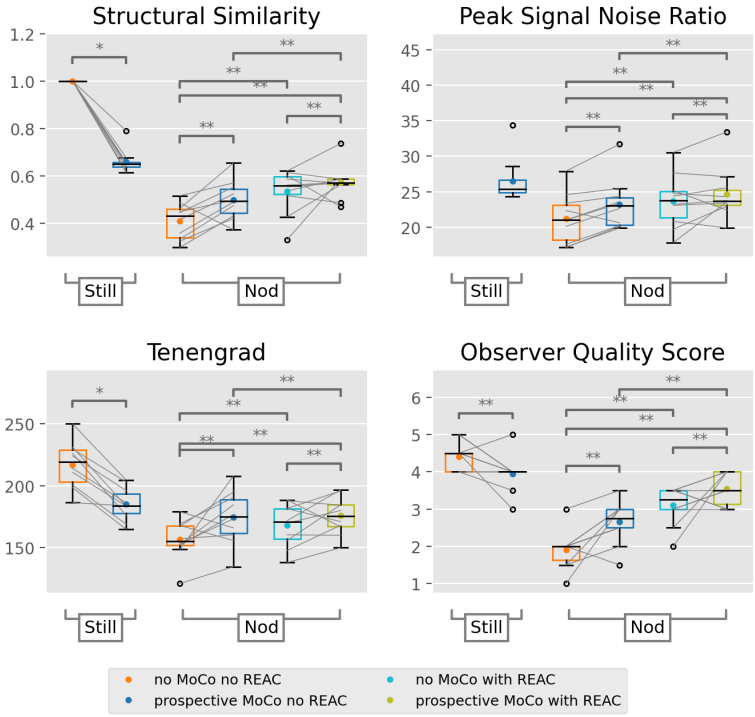


FIGURE 36: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_2 FLAIR scans with and without pMC as well as reacquisition. All metric values are markedly increased for scans with reacquisition (cyan and green). The largest improvement is observed for the combination of reacquisition and motion correction (green), which clearly outperforms the clinical standard (orange, please consult the caption of Figure 34 for figure details).

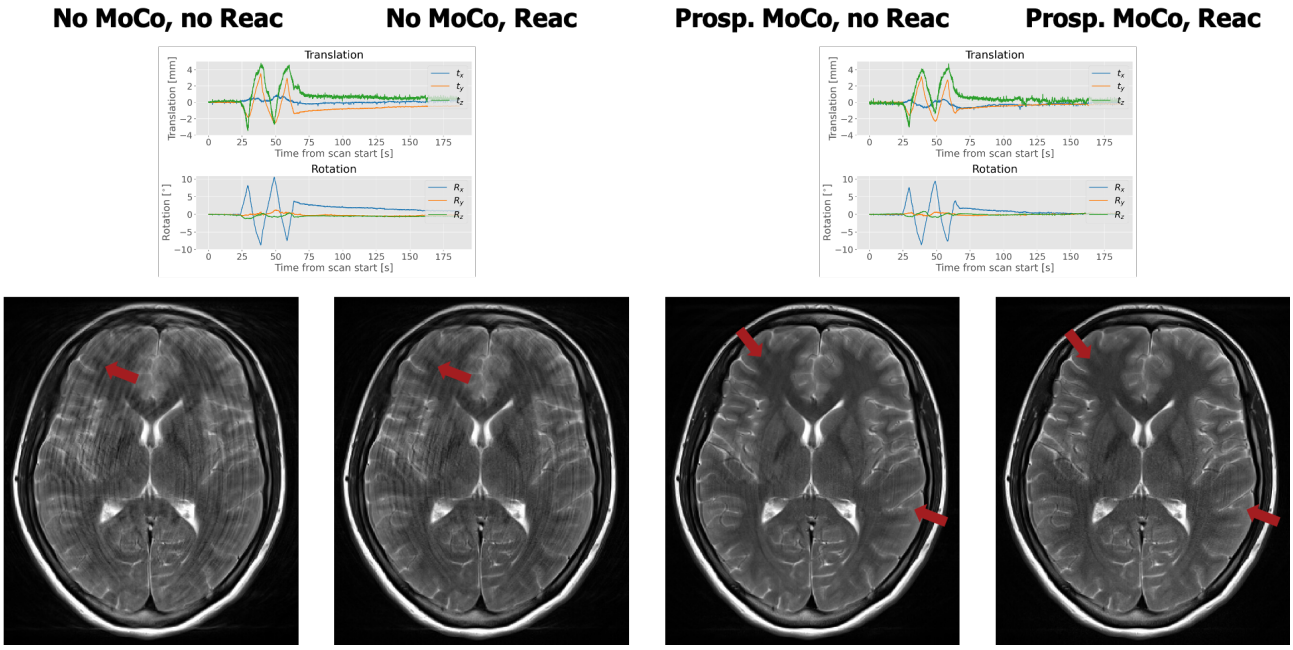


FIGURE 37: Examples of T_2 TSE images with nodding motion, comparing the reconstructed scans without and with reacquired slices for acquisitions without and with pMC. Red arrows indicate regions, where reacquisition decreased ringing as well as blurring. The differences due to reacquisition are small compared to the ones due to motion correction. They are particularly difficult to detect, when single slices are displayed next to each other. The corresponding motion curves are shown above the two respective images. All images shown were registered to the still, uncorrected scan.

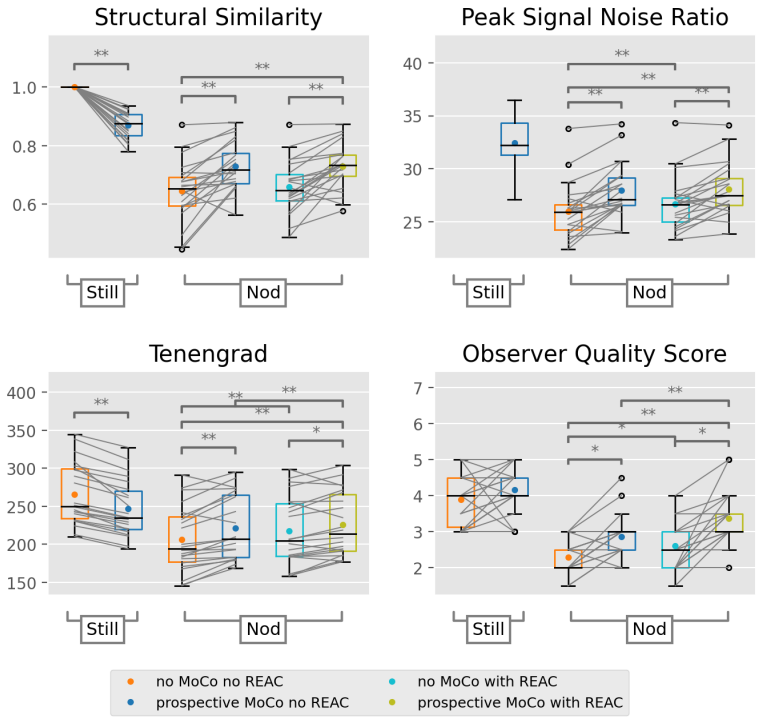


FIGURE 38: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_2 TSE scans with and without pMC as well as reacquisition. Differences between scans with and without reacquisition, but with the same motion correction setting (orange and cyan or blue and green), are comparably small (please consult the caption of Figure 34 for figure details).

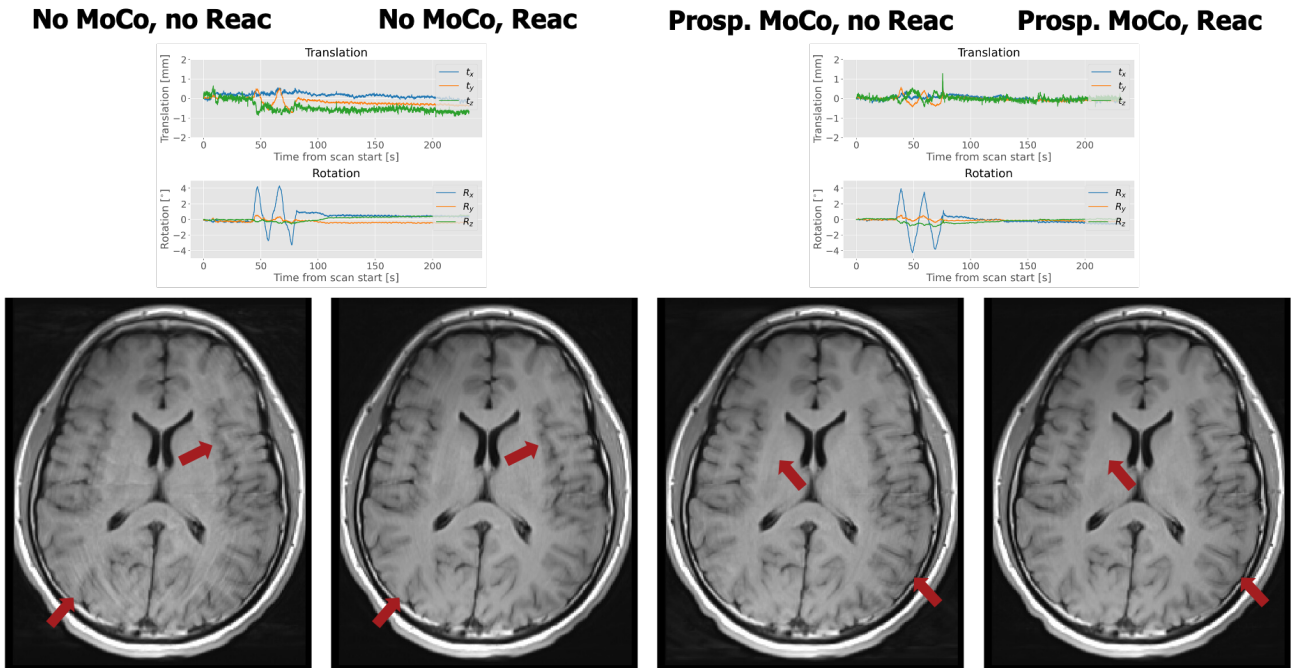


FIGURE 39: T_1 STIR image examples with nodding motion, comparing the reconstructed scans without and with reacquired slices for acquisitions without and with pMC. I highlighted areas with red arrows, where blurring and especially ringing were reduced in the scans with reacquired slices. The amount of artefacts is comparable in the scans with only reacquisition or only motion correction (second and third). The corresponding motion curves are shown above the two respective images. All images shown were registered to the still, uncorrected scan.

clinical standard (neither motion correction nor reacquisition).

Image examples for the impact of reacquisition on the T_1 STIR sequence are shown in Figure 39, which demonstrates that reacquisition decreased artefacts. For this visualisation, I chose a case where pMC led to improved image quality, more precisely images example from the same participant as in Figure 26. The analysis of image quality metrics for all subjects is shown in Figure 40. This comparison confirms the findings in the image examples, namely that reacquisition of slices with largest amount of motion increased image quality and that reacquisition had a slightly larger impact on image quality than pMC.

4.3 Cortical thickness changes

For investigating how motion and in particularly motion correction influence the estimation of cortical surfaces, I correlated cortical thickness changes with RMS displacement by vertex-wise fitting a general linear model. Like this, I compared all uncorrected and prospectively corrected scans to the still, uncorrected MPRAGE scan, as explained in section 3.5.

The analysis shown below was performed for cortical thickness estimates resulting from *FreeSurfer's* cross-sectional processing stream. Analogous analyses of estimates from the longitudinal stream can be found in the appendix (Figures B.9 and B.10).

The significance maps for scans with nodding motion in Figure 41 show that both prospective motion correction and reacquisition led to a reduction in the number of vertices with statistically significant changes, but only the combination of both completely avoided significant changes. For all scans showing significant thickness changes, Figure 42 visualises the slope of the general linear model in percent. For the majority of vertices the thickness change is around 4 %,

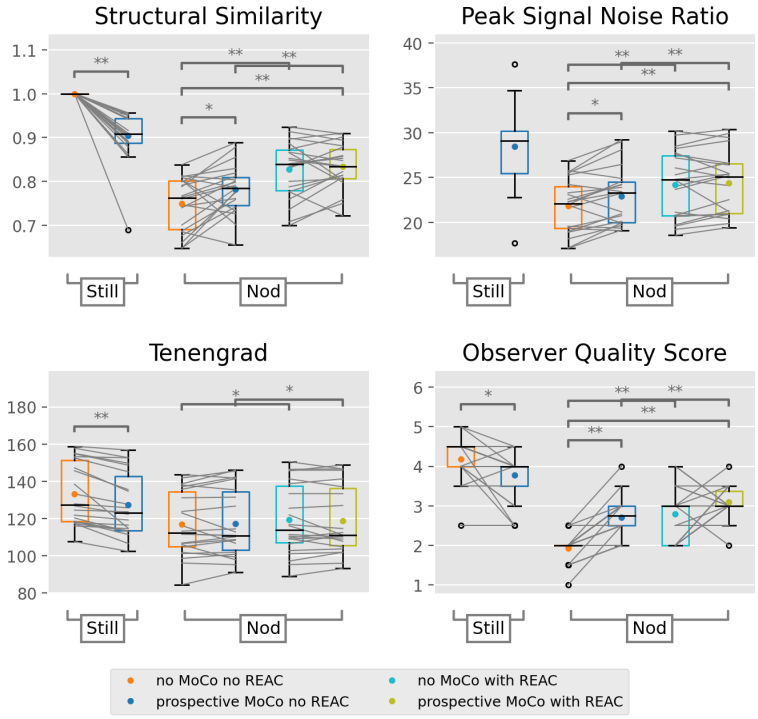


FIGURE 40: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_1 STIR scans with and without pMC as well as reacquisition. Scans with reacquisition (cyan and green) clearly obtain larger values for SSIM and PSNR, whereas the effect is smaller for observer scores and rather small for Tenengrad (please consult the caption of Figure 34 for figure details).

corresponding to approximately 0.01 mm per 1 mm increase in RMS motion. The number of vertices with larger changes is reduced in the scans with reacquisition or pMC compared to the scan without.

For the scans with shaking motion, I observed similar general patterns, i.e. the combination of pMC and reacquisition also showed no significant cortical thickness changes (please compare Figures B.11 and B.12). However, the number of significant vertices in scans with pMC and without reacquisition was slightly larger than for the scans with nodding motion.

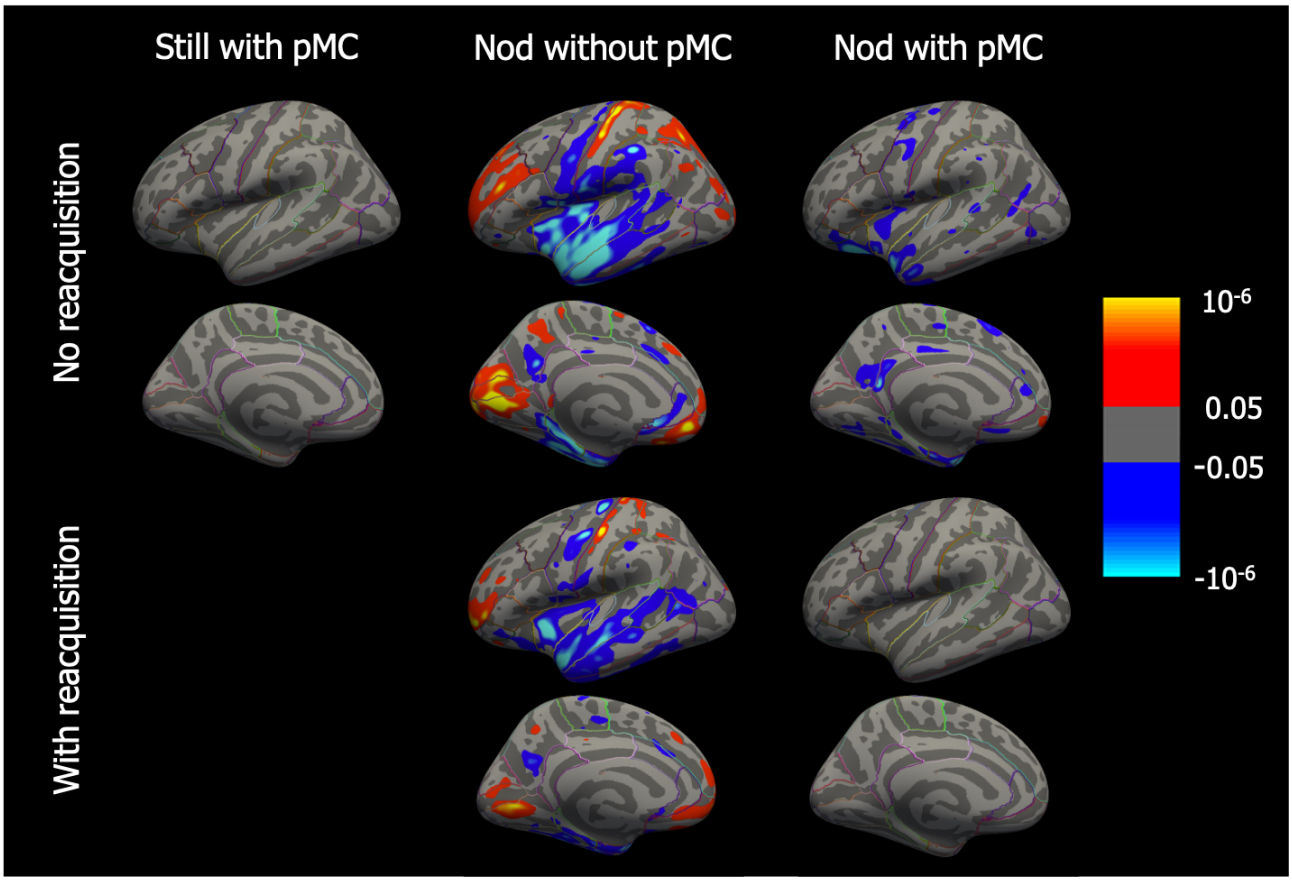


FIGURE 41: Significance maps of the left hemisphere for GLM analysis of cortical thickness changes with increased RMS displacement in nodding and motionless scans. Blue represents decreasing cortical thickness, yellow and red increasing thickness with increasing motion. The visualised p-values were FDR thresholded at `FDR=0.05` for both hemispheres. pMC did not introduce a bias on motionless scans. However, in the case of nodding motion pMC and reacquisition both decreased the number of significant vertices. The combination of both does not show any statistically significant correlation with motion.

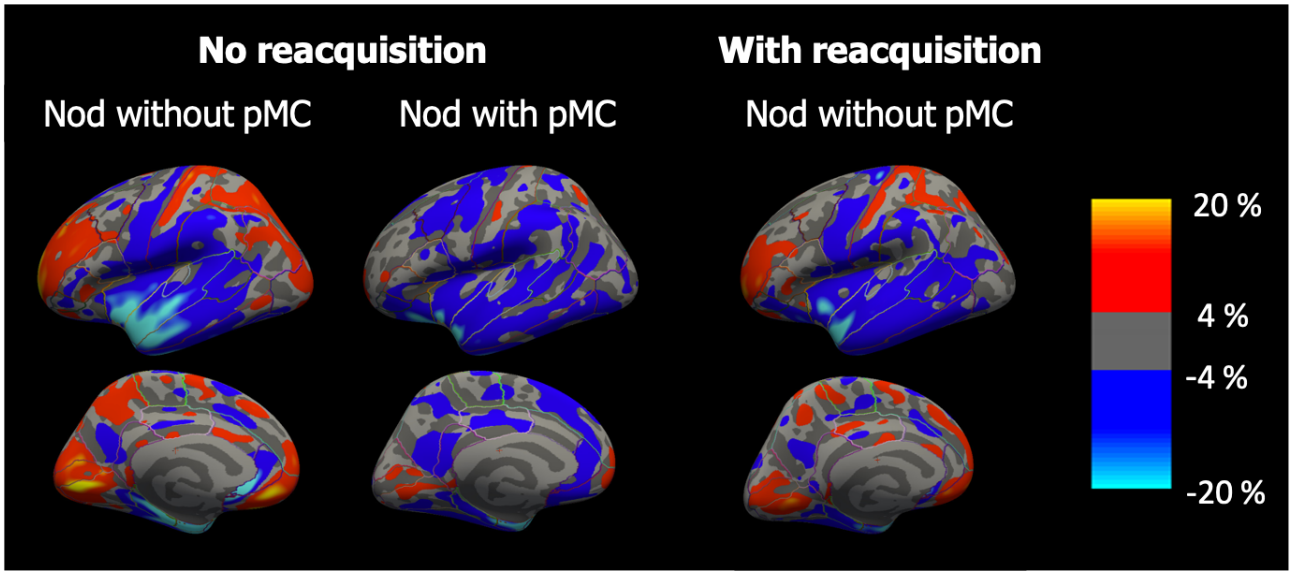


FIGURE 42: Percentage thickness changes for increased RMS displacements for scans showing a correlation in Figure 41. Blue represents a negative slope, yellow and red a positive one. The average cortical thickness over all subjects and all cortical areas was determined to 2.40 ± 0.07 mm, using the *FreeSurfer* function *mris_anatomical_stats*. Thus, a change of 4 % (dark blue or red) per 1 mm increased RMS displacement corresponds to approximately 0.10 mm thickness change, whereas a change of 20 % (light blue or yellow) corresponds to approximately 0.48 mm. The larger part of vertices for all scan types only shows very small thickness changes. The number of vertices with changes far above 4 % was decreased by reacquisition, as well as pMC.

5 Discussion

The aim of this thesis was to analyse the influence of prospective as well as retrospective motion correction on MR image quality for a full clinical children protocol. For that, we performed a healthy volunteer study with 22 subjects, which consisted of repeated scans of various sequences with different motion patterns and motion correction settings. The results of these experiments are discussed below.

5.1 Motion patterns

For our motion experiments we chose nodding as main motion pattern, based on a previous study [31, 32], where we analysed motion data during an MR examination of children with and without GA. For the no-GA-group, the dominant motion was translation along the z-axis (along the bore of the scanner) and rotation around the x-axis, corresponding to nodding movement, which is in agreement with observed motion for adults [60, 61]. This can be explained by padding of the head with cushions etc., limiting side-to-side motion, as well as especially children watching TV over the mirror or looking for their parents at their feet. For the 3D-encoded MPRAGE sequence, we additionally acquired scans with a shaking motion of slightly larger amplitude to investigate the influence of a different motion direction and amplitude.

I quantified each subject's motion per scan by calculating root mean squared, median and maximum displacement with respect to the start of each scan. These motion metrics are compared in Figures 17 and 18. We aimed for equal amounts of motion in the scans with and without motion correction of the respective type of motion, which we implemented by training the motion pattern with the volunteers before the scan as well as giving them feedback based on the motion tracking throughout the examination.

With a median of 13.1 ± 5.2 mm, (Table 6), the maximum displacement of the periodic motion within this work exaggerated the median value of 4.4 ± 3.0 mm we found for the awake children over all scans. Median displacements, on the other hand, were similar: 1.7 ± 2.7 mm compared to 2.0 ± 1.2 mm for children without GA [32]. Recall that the mentioned standard deviations represent variances of the data and not uncertainties in the context of physical variables. That is to say, the large values of standard deviations relative to the respective median values in Table 6 indicate that for this experiment the motion varied in amplitude from participant to participant with a range that is in the order of magnitude of the median of all participants.

Larger amplitudes might lead to a larger distortion of the images. However, we aimed to test the performance of the correction techniques for challenging scenarios. For the same reason, we chose to perform the motion for 40 s during acquisition of the k-space center, since motion is expected to have the largest impact on image quality in central parts of k-space. Quasi-periodic motion might lead to different artefacts than random motion patterns. Nevertheless, we considered it important that the motion of all healthy volunteers within our study was consistent and comparable, which was most easily realisable by guiding them through a periodic motion pattern.

For the T_2 TSE and T_2* sequences, Figure 17 shows significantly larger metrics values for the uncorrected compared to the corrected still acquisition. In many individual motion curves corresponding to the uncorrected still scan, I noticed some motion leading to a shift out of the original position in the beginning of acquisition, which I did not observe to the same degree in the corrected scans. This phenomenon could be explained by the participants

being startled by the louder noise of the T_2 TSE and T_2* sequences compared to the sequences acquired previously. Since the uncorrected scan was always performed first, the participants were already used to the higher noise level in the second, corrected acquisition. This effect could have been avoided by randomising the order of acquisitions in the study design. However, the difference between the motion for the two scan types was not significant, when the metrics were calculated with respect to the middle of the acquisition instead of the beginning, as Figure B.1 of the appendix shows. This indicates that the observed phenomenon should not significantly influence the results of this study.

For the nodding scenario, DWI scans showed substantially larger values of RMS and median displacements than the other sequences (compare Figure 18). The values for maximum displacements were in the same range as for the other sequences, with slightly larger mean values. These could be explained by the fact that the duration of a DWI scan was only 42 s, so the volunteers might have been under pressure due to moving throughout the whole scan and not only for a part of the scan, as they were during the previous sequences. The differences in RMS and median displacements, on the other hand, are explained by the fact that these metrics were calculated over the whole scan period and DWI scans were significantly shorter than scans of the other sequences. The level of motion between DWI scans with and without motion correction was similar, which is most important, since I only compared image quality metrics within a sequence.

5.2 Image quality measures

Quantitatively assessing image quality is challenging, especially for medical imaging, since no single image quality measure is sensitive to all possible artefacts and thus, fully correlates with radiological evaluation [27, 62, 63]. In this work, I chose to use three different image quality metrics that are commonly used for quality assessment in medical imaging [26, 35, 64]. Peak signal-to-noise ratio is based on the mean squared error of the image relative to a reference image, thus analysing pixel-wise differences or, as the name suggests, noise between the image and the ground truth. Structural similarity index, on the other hand, was developed to better represent perceived image quality by taking structural information into account [51]. It compares patterns of pixel intensities by computing covariances between image patches. Even though PSNR and SSIM respond to a different extent to different image distortions, they are not fully independent metrics [65]. Which is why I additionally used the Tenengrad measure, which as a gradient measure is sensitive to the sharpness and contrast differences of edges or tissue boundaries [62]. Yet, I observed that neither of the three image quality metrics is sensitive to mild to moderate prevalence of ringing artefacts. This is because all metrics average over individual pixels and thus, smooth out the thin lines of intensity changes defined as ringing. Large ringing is picked up by all metrics, though.

PSNR and SSIM are full-reference metrics, which means that they are always calculated relative to a ground truth image, which within this work is the still, uncorrected scan. Thus, these metrics are not appropriate for comparing the motionless scans and were only shown for completeness. In order to get an intuition of what value these reference metrics show for a quasi-optimal scenario, I calculated values for two sets of repeated still and uncorrected MPRAGE and TSE scans in Table 5. The largest values were 32.07 and 0.86 for PSNR and SSIM of a MPRAGE test-retest set. These values should be kept in mind, when comparing metrics for the different motion types and motion correction settings.

I would additionally like to highlight the effect of image registration and hence, resampling

of voxel intensities. Intensity smoothing due to interpolation could be wrongly identified as motion artefact. Even though all scans are transformed into the space of the still, uncorrected scan, the effect might be larger on the scans with movement, since they can be expected to be less aligned with the ground truth scan. Furthermore, I would expect this effect to be especially relevant for the 2D-encoded sequences with slice thicknesses of 4 or 5 mm. This could explain the lower metric values for TSE test-retest experiments compared to the values for the isotropically sampled MPRAGE in Table 5.

Next to the quantitative metrics, I analysed observer quality scores given by two radiography students, which are based on Table 3 and evaluate diagnostic image quality. Please note that clinical usability of medical images does not always correspond to artefact-free images with respect to research standards. However, we chose this definition, since the main objective of the healthy volunteer study was to evaluate the performance of motion correction in a clinical setting.

Additionally, I would like to highlight that the comparison of image quality scores by radiographers for retrospectively corrected scans with all other scans is not completely fair, since retrospectively corrected scans have been processed with a different pipeline, as described in section 3.1. Thus, quality scores might be in favour of scans reconstructed with the more elaborate pipeline at the scanner. This could be the reason why retrospectively corrected still MPRAGE scans in Figure 21 do not achieve as high values as uncorrected or prospectively corrected scans. However, I chose to not use all scans reconstructed with the retrospective pipeline, since I detected some severe ghost-like artefacts in some of the prospectively corrected scans, when they were reconstructed with the off-line reconstruction pipeline (compare Figure A.1 of the appendix). These artefacts could have been judged as motion artefacts and thus, biased the quality scores with a disadvantage for the scans with pMC. This trade-off can only be solved by fixing the retrospective pipeline, which remains future work.

In the following, I will discuss the performance of prospective and retrospective motion correction in terms of image quality measures sequence by sequence.

T_1-weighted MPRAGE Both motion correction techniques considerably improved image quality for scans with nodding and shaking motion by reducing blurring and enhancing contrast between white and grey matter, as illustrated in example images (Figures 20 and B.3). When analysing those example images, it is important to keep in mind that only one exemplary slice is shown and consequently not all differences and details can be appreciated. However, those example images are suitable for developing a first impression of the effects of the different correction techniques.

The comparison of image quality measures in Figures 21 confirmed the observation, i.e. it showed that image quality significantly increased from scans without MoCo over scans with retrospective MoCo to scans with prospective MoCo. These results are in agreement with theoretical expectations, since rMC is limited amongst others by the phenomenon of sparsely sampled regions after correcting the raw data for rotations, as illustrated in Figure 5.

Furthermore, I observed higher values of all measures for the scans with nodding compared to shaking motion. On the one hand, this could be explained by the larger motion amplitude as well as average motion connected to the shaking in our experiment, as analysed in Figure 18. Since a larger amount of motion leads to a a larger amount of artefacts, the correction of these could be more complicated. On the other hand, the direction of motion relative to the sequence's orientation and phase encoding direction (Table 4) needs to be taken into account. In contrast to nodding, shaking motion is a through-plane-motion. Since the MPRAGE sequence

is 3D-encoded, this means that the main direction for shaking of the head is along one of the phase encoding directions (R → L), whereas the main motion direction in the case of nodding corresponds to the frequency encoding direction (S → I). Motion during and in the direction of a field gradient leads to phase accumulation and thus, in particularly falsifies phase encoding of the signal [12].

For three prospectively corrected scans with shaking motion and one with nodding motion I observed severe ringing artefacts, especially in the cases without reacquisition. Some of these outliers are visible with clearly smaller metric values and observer quality scores in Figure 21. For those scans I noticed additional peaks in the motion curves before and after the shaking motion, where the participants were instructed to lie still. An exemplar image for one of those shaking scans and the corresponding motion curve can be consulted in Figure C.1 of the appendix. A replay of the saved point clouds showed that the volunteers were not moving their head, but were strongly blinking or deforming their face in other non-rigid manners and thus, falsifying the tracking estimates. These erroneous motion estimates then propagated into correction of a non-existing motion, which caused the observed ringing. TracInnovations collaborated with us to understand this problem and is working on a solution to detect and filter those non-rigid deformations in real-time.

The image quality measures for scans with and without selective reacquisition, compared in Figure 34, show a clear trend. Reacquisition led to an additional increase in image quality. Metric values for scans with only pMC and for scans with only reacquisition were on similar levels, with a small advantage for pMC. However, when comparing these results, one should keep in mind that for the MPRAGE sequence, we reacquired 16 T_R's, corresponding to 32 s of reacquisition, which is almost the whole 40 s motion period. In a real-case scenario, I would expect the children's motion to be more distributed over the whole scan time and I would not expect the children to lie perfectly still during the reacquisition period. Thus, reacquisition might not lead to the same degree of improvement as in these healthy volunteer experiments. On the other hand, for cooperative children performing only a small amount of motion, it might not even be necessary to reacquire 16 T_R's. For such cases it would be beneficial to only reacquire slices where the motion exceeded a certain threshold and thus, being able to reduce the total scan time. However, a proper way to define this threshold requires future research.

These findings fit well into the literature, where retrospective and in particularly prospective motion correction for the MPRAGE sequence have been extensively reported to be effective, using markerless as well as navigator based motion tracking with and without reacquisition [13, 14, 24, 66].

T_2-weighted FLAIR For the T_2 FLAIR example images in Figure 22 it is not clear without doubt whether prospective or retrospective motion correction led to higher image quality. Blurring was more reduced in scans with pMC, whereas those show more ringing artefacts than the ones with rMC. The latter could be caused by noisy motion estimates, which are easier to filter retrospectively. However, these observations are based on a single slice for one volunteer. In the following I will discuss the performance of motion correction in more detail based on the image quality measures. The comparison in Figure 23 indicates that pMC led to increased image quality in the nodding scans. Even though the calculated metrics are not always sensitive to a mild degree of ringing, those artefacts were considered for quality scoring, if they impede a diagnosis. Both Tenengrad and observer quality scores show a decrease in image quality for the still scan with pMC compared to the uncorrected as well as the scan with rMC. In agreement, I observed slightly more blurring in the still scans with

pMC (compare an example in Figure B.4 in the appendix), even though interpolation related smoothing complicates comparisons of registered images. This slightly reduced image quality for still scans with pMC could again be an indication that the motion tracking is noisy and more filtering needs to be applied. However, real-time filtering of the motion tracking curves is challenging and needs to be investigated in future studies. Nevertheless, the reduction in image quality was small and still scans with pMC resulted in a median quality score of 4, which is fully diagnostic.

SSIM, PSNR and observer quality scores did not reach as high values for the prospectively corrected images with nodding motion as the MPRAGE scans, for instance a median quality score of 2.75 for FLAIR compared to a median of 3.5 for MPRAGE scans. Since orientation and phase encoding directions were the same for those two sequences, the origin for this is not connected to the direction of motion relative to spatial encoding directions. However, the difference could be due to the fact that for the MPRAGE sequence within-echo-train correction was performed, whereas the FLAIR scans were only corrected before each echo train, since Frost et al. [13] showed that within-echo-train correction was not as effective for the T_2 SPACE sequence. The latter seemed to be more sensitive to continuous motion, which might be because it is not spoiled and thus, has more complicated signal dependencies [13]. The investigated T_2 SPACE is similar to the FLAIR sequence used within this work, only without inversion as magnetisation preparation.

Reacquisition led to an additional increase in image quality, indeed a larger increase was observed by only reacquisition than by only pMC. Combined, the image quality was improved considerably compared to the clinical standard. The median quality score was with 3.5 again not as high as a median of 4 for the MPRAGE sequence, but no scan was scored less than 3, which only has to be repeated if an abnormality was detected. Please note that similar to the comment for the MPRAGE sequence 7 T_R's, corresponding to a duration of 35 s, were reacquired and results cannot necessarily be expected to be as good for scans of children.

T_2-weighted TSE For the T_2 TSE sequence mild motion artefacts, in particularly ringing, could already be detected in the still, uncorrected scans, as shown exemplary in Figure 24. Observer quality scores varied between 3 and 5 for the still, uncorrected TSE (compare Figure 25), while none were below 4 for MPRAGE or FLAIR. This can be explained by the TSE being a 2D-encoded sequence in contrast to the 3D-encoded sequences discussed above. The slice-selective excitation makes 2D-encoded sequences more sensitive even to small amounts of motion due to the spin history effect (section 2.2). I detected a reduction of ringing due to prospective motion correction in many still scans, even though this trend was not visible in the observer scores based on diagnostic image quality. It is not possible to retrospectively correct artefacts caused by the spin history effect for 2D-encoded sequences, thus we decided to only evaluate rMC for the 3D-encoded sequences, where it can demonstrate it strengths.

In the nodding scenario, prospective motion correction led to a significant increase in all image quality measures, even though there were a few more outlier cases, where pMC did not lead to an improvement or where residual ringing could not be avoided by pMC. I examined the motion data for those cases, but did not find any systematic anomalies as e.g. increased tracking noise or erroneous tracking data. It remains future work to find a system behind this infrequent occurrence of residual ringing in corrected nodding scans. However, I often observed decreased blurring of underlying structures, even if ringing artefacts could not be completely eliminated and the radiologists are used to ignore a certain degree of ringing in motion sensitive 2D-sequences.

The effects of reacquisition in Figure 38 were rather small, especially for quantitative metrics. One reason for the reacquisition not leading to an as large improvement as observed for the other sequences might be the fact that for the T_2 TSE sequence the k-space is sampled twice and then averaged. The volunteers only moved during the first sampling. A small mismatch between the position of the slice in the first acquisition and in the reacquisition (due to e.g. noise in the tracking data) might lead to inconsistencies when averaging with the corresponding slice from the second sampling.

To my knowledge, prospective motion correction has not been evaluated in detail neither for the 2D-encoded TSE, nor for the STIR sequence in the literature, even though [67] display image examples for one subject.

T_1-weighted STIR The results for the second 2D-encoded sequence, the T_1-weighted STIR, were less conclusive. Even though median observer scores for prospectively corrected T_2 TSE and T_1 STIR scans with nodding motion were comparable (3.0 compared to 2.75), I visually observed more cases where pMC did not lead to improved image quality or a considerable amount of ringing was present in the corrected STIR scans. This was also reflected in the image quality metrics in Figure 27, where more outlier scans could be observed with a decreasing value from uncorrected to corrected scans. In general, motion artefacts were only present in every second slice, since we were acquiring two concatenations and the participants were only instructed to move during the first concatenation.

For many of the scans with residual amount of ringing I observed a shift of the head out of its original position after the nodding motion, even though we marked the reference position with masking tape on the inside of the scanner bore. For some of these scans with pMC, ringing was even apparent in every slice, not only in slices corresponding to the first concatenation. The reason for that might be changed B_0 inhomogeneities, induced by shifting the head out of the original position.

An explanation why pMC led to more successful cases for the TSE than for the STIR is the additional inversion pulse and delay time of 1600 ms between inversion and readout, which is not present in the TSE. If inverted and imaged slice do not perfectly match due to e.g. tracking noise, inconsistencies could be caused in the acquired signal. In future studies it would be interesting to investigate whether BLADE or PROPELLER acquisition with a rotating k-space trajectory lead to more robust results for the STIR sequence. This type of acquisition oversamples the k-space center and thus, inconsistent data points could be discarded [68].

Reacquisition had a marked effect on image quality for the STIR. The values of image quality measures were similar, if not even higher for scans with only reacquisition compared to scans with only pMC. Again, I would like to point out that we were reacquiring 10 T_R's, corresponding to a duration of 41 s, which is equivalent to the whole nodding period. Results due to reacquisition cannot be expected to be as promising for all children.

Furthermore, the information content of the T_1-weighted STIR is very similar to the one of the high-resolution MPRAGE, for which motion correction led to a considerable and robust increase in image quality. Both sequences are used for examining e.g. bleedings and migration defects. Consequently, I would recommend to mainly focus on acquiring a good MPRAGE scan and only optionally acquire the STIR, when it is needed for a specific diagnosis in addition to the MPRAGE.

T_2*-weighted GRE For the T_2*-weighted sequence, no improvements could be observed for prospective motion correction of nodding scans. Since this observation was already obvious

in the quantitative image quality analysis (Figure 29) and by looking through the images (examples in Figure 28), observer quality scoring was disregarded.

The T_2*-weighted, GRE-based sequence is designed to be sensitive to magnetic field inhomogeneities caused by susceptibility differences near tissue boundaries, since T_2* is defined as a combination of T_2 relaxation and B_0 inhomogeneities (section 2.1.3, [17]). However, a moving object induces changes in magnetic field inhomogeneities which cannot be corrected by adjusting the position of the FOV. Thus, the field correction, done by shimming before acquisition, seems to move in opposite direction as the object. Especially, nodding and shaking rotations, whose axes are perpendicular to B_0, lead to significant field inhomogeneity changes. These inhomogeneity changes could only be avoided by dynamic shimming, which is still an ongoing research question and has not yet been combined with pMC [12, 22].

An alternative approach to avoid these difficulties with pMC of T_2* GRE-based sequences would be to acquire a shortened 3D-encoded, susceptibility weighted, EPI-based sequence with an acquisition time of 52 seconds. Due to the short acquisition time, the motion can be monitored during the acquisition and the scan repeated, if a substantial amount of motion is noticed. We acquired two prototype scans, where we compared still, uncorrected scans of this shortened EPI-based SWI sequence with the above discussed T_2*, as well as a clinical standard GRE-based SWI sequence. One set of these prototype scans is compared in Figure C.2 of the appendix. The images show similar content and collaborating radiologists judged the accelerated sequence as a sufficient substitute for clinical SWI or T_2* sequences. In Figure C.3, I compare acquisitions of the 50 s SWI sequence with and without pMC for still scans as well as scans with a shift of the head, simulating a small motion. pMC seems to slightly reduce blurring and signal loss.

Diffusion-weighted sequence DWI sequences are in general considered to be relatively robust towards slow motion, as present in our experiments, whereas severe signal dropouts are observed for strong and fast motion [22]. Nevertheless, the nodding scan without pMC in Figure 30 showed the presence of small motion artefacts, such as blurring in the TRACE-weighted scans and signal variations in ADC maps. Prospective motion correction was able to reduce these artefacts. Differences, especially in the ADC maps were difficult to notice in individual scans, I could only identify them with certainty when comparing them with the other scans of the same volunteer. Thus, it would not have been possible to create meaningful observer scores in a fully blinded fashion and we decided to not perform observer quality scoring for this sequence.

However, the quantitative image quality metrics I calculated for the TRACE-weighted images in Figure 31, confirm my observations on the example images, namely that prospective motion correction led to a significant increase in image quality. For the ADC maps, representing real physical constants, I did not calculate image quality metrics, but computed difference images to the ground truth scan and compared mean absolute differences (Figure 32). I did find a significant decrease, corresponding to an increase of image quality, for the prospectively corrected scan. Analogously to the analysis of Berglund et al. [54], I also calculated mean and standard deviation of the histograms of the individual difference images. Berglund et al. showed a decrease of both measures due to pMC for a single subject performing four different motion patterns, including nodding motion. In our case, a statistically significant decrease can be observed for standard deviations, which is in accordance with the significant decrease in mean absolute differences. The means of the histograms, however, did not differ significantly, pointing to the fact that the positive bias of approximately 20 $\frac{\mu m^2}{s}$ could not be corrected by

pMC. Nonetheless, this value is relatively small compared to the mean ADC value for white matter of 700 $\frac{\mu m^2}{s}$ [17].

In the following paragraphs, I would like to summarize a few thoughts on the ringing artefact, which occurs frequently in our data set and originates from various sources, not all of which I could clearly identify in the above discussion. The appearance or reasons for ringing might overlap with appearance of ghosting artefacts, as ringing is often observed to be parallel to the skull.

First of all, ringing was present in many uncorrected nodding and some uncorrected still scans. In these cases, it was a clear motion artefact originating from signal inconsistencies in k-space, which themselves can be caused for instance by the spin history effect.

A possible reason for ringing in corrected scans is noisy or erroneous tracking data. I observed the latter for a few MPRAGE scans with shaking motion, where the subjects were deforming their face in a non-rigid manner by e.g. squinting their eyes or moving their nose. Noise in the tracking data, on the other hand, can result from different phenomena, as for example vibrations of the camera arm due to fast switching of gradients. It can also be due to the subject's head moving out of the reference position, which makes the overlap of the tracked point cloud with the reference point cloud smaller and leads to a higher noise level in the estimation of tracking data. Maclaren et al. [22] address this topic by showing an example, where noisy motion data leads to a severe amount of ringing. Tracking noise can be reduced by updating the reference point cloud during the scan session in order to maximise the overlap throughout the whole session. Furthermore, more sophisticated filtering approaches than the median filter we were using in this study could be applied to the tracking data before performing prospective motion correction. However, there is a trade-off between the additional latency of motion estimates induced by heavier filtering and the reduction of noise. Thus, further studies are required to find the optimal filtering solution.

Ringing artefacts in MR images can in general also be caused by the Gibbs phenomenon. The origin of these special ringing artefacts lies in the truncation of the Fourier series, since only a finite number of frequencies are sampled. Smooth signal changes can be well represented by a small number of terms in the Fourier series, whereas sharp intensity changes cause overshoot oscillations [69]. However, Gibbs ringing artefacts are not typical motion artefacts, they rather depend on the number of phase encoding steps relative to the size of the field of view.

5.3 Cortical thickness changes

In the above section about image quality measures, I discussed the performance and the relevancy of motion correction and reacquisition for diagnostics in the clinical setting. In the following paragraph I will additionally review the relevancy of these techniques in research applications. Reuter et al. [4] demonstrated that motion can systematically bias the derivation of morphometric estimates, such as cortical thickness, which is especially concerning for diseases that come along with an increased tendency for motion, such as epilepsy, Alzheimer's or Huntington's disease. The authors recommend affected studies to restrict movement and to control for motion in the statistical model. In this work, I investigated whether motion correction and reacquisition can reduce the observed bias. For this purpose, I correlated cortical thickness changes with root-mean squared motion comparing all motion types and motion correction settings individually with the ground truth scans.

I performed the main analysis on MPRAGE images processed with *FreeSurfer's* cross-

sectional stream, because I don't assume the expected application to necessarily acquire paired data. At the Neurobiology Research Unit, for instance, mainly cross-sectional studies are performed, which compare a group of diseased patients (e.g. suffering from major depressive disorder) with a healthy control group. However, should repeated scans of the same subject for different time points be available, as e.g. for longitudinal Alzheimer studies, it would be beneficial to process them in *FreeSurfer's* longitudinal stream [70], which first estimates surfaces on a within-subject template and then fine-tunes the estimates for the individual scans. Thus, data processed in the longitudinal stream can be expected to be less corrupted by motion. I appended an analogous analysis for longitudinally processed data (Figures B.9 and B.10), which confirms the expectation of being more robust towards motion. However, the same trends are observed as discussed below for the cross-sectional stream.

The analysis was only performed for uncorrected and prospectively corrected scans, since I only recently had retrospectively corrected data available with bias field correction and thus, I did not have enough time to process all scans in *FreeSurfer*. Expanding the analysis with retrospectively corrected data remains future work.

The significance maps and percentage thickness changes in Figures 41 and 42 showed that the bias in estimating cortical thickness induced by nodding motion could be considerably reduced by pMC as well as reacquisition. The combination of both eliminated any significant thickness changes. This analysis is in accordance with the results from the image quality measures and indicates the relevancy of motion correction in research studies on movement disorders or diseases with an increased tendency for motion. However, I would like to refer to the above comments I made on reacquisition when discussing image quality measures for the MPRAGE sequence. Depending on the motion pattern, reacquisition might not be as powerful in a real-case scenario as in our experiments.

Furthermore, analogous analyses of scans with shaking motion in Figures B.11 and B.12 of the appendix showed that pMC did not reduce the bias in thickness estimates to the same degree as in nodding scans, which I have already observed for the image quality measures. It can be explained with the larger motion amplitude of the shaking motion as well as with the different motion direction relative to the orientation of phase encoding gradients.

Direct comparisons of our results for scans without motion correction with the results of Reuter et al. [4] are difficult, since they acquired five scans per subject (two still scans, as well as scans with nodding, shaking and free motion) and performed the analysis with a linear mixed effects model. I decided to use a general linear model with a smaller number of parameters for my analysis, since we only acquired three scans and are especially missing the one with free motion, having a considerably larger RMS motion (compare Figure 1 in [4]).

A similar analysis, using a GLM to evaluate the performance of motion correction in reducing bias, was performed by Frost et al. [13]. They compared cortical thickness changes for scans without pMC and scans with before- as well as within-echo-train pMC for shaking motion with different velocities. In agreement with our results, motion correction reduced the number of significant changes. They even observed a complete elimination of significant changes for the scans with within-echo-train correction. However, their analysis was only performed for one subject, which underlines the strengths of this work, namely confirming parts of the results from Frost et al. [13] with a larger sample size.

5.4 Limitations

The presented study is not without limitations. Firstly, our participant population only consisted of healthy adults, which means that I can merely draw conclusions about overall image quality, but I cannot make statements about how motion correction influences the diagnosis of certain pathologies like tumors or lesions. Since prospective correction "only" updates the scanner's field of view, I would not expect the risk of removing indications for certain pathologies to be as high as for instance for deep-learning based retrospective techniques [63, 71]. However, this topic remains speculative and a particular disadvantage of prospective motion correction is its irreversibility, since the correction is directly applied during acquisition. To gain additional insight into its influence on pathologies, we are going to evaluate the performance of prospective motion correction on a clinical child population in a follow-up study included in the MoCo project at the Neurobiological Research Unit.

Secondly, we were not able to acquire all sequences of the protocol for 20 subjects, which was the sample size I estimated to be required for sufficient statistical power on prototype scans of the MPRAGE sequence in section 3.6. As described in section 3.2, sequence parameters and motion correction implementation needed to be changed for the FLAIR and DWI sequence after the twelfth volunteer. We decided to scan at least 10 volunteers with those sequences in order to see some trends in image quality measures, when comparing the different types of motion and motion correction settings. Thus, 22 sets of scans were available for MPRAGE, STIR and TSE, whereas I was only able to analyse 10 sets of scans for FLAIR and DWI, which reduces the statistical power for comparisons of those sequences. With the T_2* sequence we only scanned 19 volunteers, since the results were rather clear and we used the last three volunteers for prototype scans of the EPI-based SWI sequence.

Thirdly, the results presented in this work are limited by the fact that for all sequences except the MPRAGE, only one motion pattern was performed. Nodding motion has been reported to be the most prevalent motion in the MR scanner [31, 32, 60]. Nevertheless, it is known that both motion artefacts and presumably also the effectiveness of motion correction depend on the direction of motion relative to the orientation of the scan [12]. However, within this healthy volunteer study it was not practicable to include several motion patterns for all sequences, since the duration of the scan sessions already was a little more than 1.5 hours. It is difficult to stay focused, when lying in the scanner for such a long time, and we already observed decreasing concentration for some volunteers towards the end of the session.

Furthermore, the image quality metrics SSIM and PSNR are not applicable for analysing differences in still scans, due to the fact that they are calculated relative to the still, uncorrected scan, as discussed above. This problem could have been avoided by acquiring a second still scan without MoCo for each sequence, as independent ground truth, to which all remaining scans are compared, including the other still, uncorrected scan. However, as I already mentioned in the previous paragraph, our total scan time was already quite long, which is why we decided to not acquire a second ground truth scan.

6 Conclusion

One of the most challenging reasons for artefacts in the field of Magnetic Resonance Imaging is head motion, which is particularly difficult to avoid when young children are scanned. This thesis set out to evaluate the performance of prospective as well as retrospective motion correction techniques on a full clinical pediatric protocol consisting of two 3D- and four 2D-encoded sequences. For that, we acquired scans with varying motion patterns and motion correction techniques, including reacquisition of slices with heavy motion, in a healthy volunteer study with 22 participants. I compared image quality metrics and observer quality scores to investigate the clinical relevance of motion correction and selective reacquisition. In addition, I correlated cortical thickness estimates with motion in order to evaluate the performance of the proposed techniques in research applications.

The results showed that the combination of prospective motion correction and reacquisition led to highest image quality for the T_1 MPRAGE, the T_2 FLAIR, the T_2 TSE and to some extent the T_1 STIR, with considerable improvements compared to the clinical standard without motion correction and without reacquisition. Since STIR scans yield almost the same information as the high-resolution MPRAGE scans, I recommend to focus on acquiring a high-quality MPRAGE scan and only acquire the STIR, if there is a specific clinical indication for it. The DWI sequence proved to be relatively robust to slow motion, but showed a slightly improved image quality due to prospective motion correction. The T_2* sequence, on the other hand, showed no increased image quality due to motion correction, which is why I suggest to substitute it with a fast EPI-based SWI scan, which can be easily repeated, if a large amount of motion is observed. Cortical thickness analyses showed that prospective motion correction and reacquisition have the potential to reduce motion related biases in morphological estimates.

All in all, this thesis confirmed that prospective motion correction and reacquisition significantly improved image quality compared to clinical standard acquisitions. It indicates that the potential risks of GA in young children could be avoided and clinical routines could be accelerated, which currently are considerably slowed down by the complicated set up connected to GA. Additionally, the findings of this thesis are a fundamental basis for setting up the protocol of a follow-up study, in which the influence of motion correction as well as training and preparation is investigated for a clinical pediatric population. Table 7 demonstrates that the acquisition times of the proposed protocol are comparable with the clinical standard and the whole protocol, including all optional scans, can be acquired in less than 20 minutes.

Table 7: Comparison of acquisition times for the different sequences in the original clinical MR protocol as well as the protocol suggested with the findings of this work. Times MPRAGE, FLAIR, TSE and STIR sequences are with reacquisition of a set amount of slices, corresponding to 32 to 41 s. Sequences in italics instead of bold letters are optional, depending on the indication of the scanning.

	Original protocol	**Suggested protocol**
T_1 MPRAGE	4 m 26 s	5 m 12 s
T_2 FLAIR	2 m 44 s	4 m 47 s
T_2 TSE	2 m 27 s	3 m 6 s
DWI	2 m 7 s	42 s
SWI	4 m 54 s	52 s
T_1 STIR	3 m 18 s	3 m 51 s
Total time	**19 m 56 s**	**18 m 30 s**

List of Figures

List of Tables

References

[1] R. W. Brown. *Magnetic Resonance Imaging: Physical Principles and Sequence Design: Second Edition*. Wiley Blackwell, June 2014. ISBN: 9781118633953. DOI: 10.1002/9781118633953.

[2] Onur Afacan et al. "Evaluation of motion and its effect on brain magnetic resonance image quality in children". In: *Pediatric Radiology* 46.12 (2016), pp. 1728–1735. ISSN: 14321998. DOI: 10.1007/s00247-016-3677-9.

[3] Erna Törnqvist, Åsa Månsson, and Inger Hallström. "Children having magnetic resonance imaging: A preparatory storybook and audio/visual media are preferable to anesthesia or deep sedation". In: *Journal of Child Health Care* 19.3 (2015), pp. 359–369. ISSN: 17412889. DOI: 10.1177/1367493513518374.

[4] Martin Reuter et al. "Head Motion during MRI Acquisition Reduces Gray Matter Volume and Thickness Estimates". In: *Neuroimage* 107 (2015), pp. 107–115. DOI: 10.1016/j.neuroimage.2014.12.006.

[5] Alysha Gilmore, Nicholas Buser, and Jamie L. Hanson. "Variations in Structural MRI Quality Impact Measures of Brain Anatomy: Relations with Age and Other Sociodemographic Variables". In: *Biorxiv* (2019), p. 581876.

[6] Jalal B. Andre et al. "Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations". In: *Journal of the American College of Radiology* 12.7 (2015), pp. 689–695. ISSN: 1558349X. DOI: 10.1016/j.jacr.2015.03.007. URL: http://dx.doi.org/10.1016/j.jacr.2015.03.007.

[7] Henrica M.A. De Bie et al. "Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans". In: *European Journal of Pediatrics* 169.9 (2010), pp. 1079–1085. ISSN: 14321076. DOI: 10.1007/s00431-010-1181-z.

[8] Stine B. Runge et al. "Children centered care: Minimizing the need for anesthesia with a multi-faceted concept for MRI in children aged 4–6". In: *European Journal of Radiology* 107.July (2018), pp. 183–187. ISSN: 18727727. DOI: 10.1016/j.ejrad.2018.08.026. URL: https://doi.org/10.1016/j.ejrad.2018.08.026.

[9] Jakob M. Slipsager et al. "Quantifying the Financial Savings of Motion Correction in Brain MRI: A Model-Based Estimate of the Costs Arising From Patient Head Motion and Potential Savings From Implementation of Motion Correction". In: *Journal of Magnetic Resonance Imaging* 52.3 (2020), pp. 731–738. ISSN: 15222586. DOI: 10.1002/jmri.27112.

[10] *FDA Drug Safety Communication: FDA review results in new warnings about using general anesthetics and sedation drugs in young children and pregnant women.* https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-fda-review-results-new-warnings-about-using-general-anesthetics-and. Issued: 12-14-2016.

[11] Jeana E. Havidich et al. "Preterm versus term children: Analysis of sedation/anesthesia adverse events and longitudinal risk". In: *Pediatrics* 137.3 (2016). ISSN: 10984275. DOI: 10.1542/peds.2015-0463.

[12] F. Godenschweger et al. "Motion correction in MRI of the brain". In: *Physics in Medicine and Biology* 61.5 (2016), R32–R56. ISSN: 13616560. DOI: 10.1088/0031-9155/61/5/R32.

[13] Robert Frost et al. "Markerless high-frequency prospective motion correction for neuroanatomical MRI". In: *Magnetic Resonance in Medicine* 82.1 (2019), pp. 126–144. ISSN: 15222594. DOI: 10.1002/mrm.27705.

[14] M. Dylan Tisdall et al. "Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI". In: *Magnetic Resonance in Medicine* 68.2 (2012), pp. 389–399. ISSN: 07403194. DOI: 10.1002/mrm.23228.

[15] Jakob M. Slipsager et al. "Clinical Implementation of Fast Markerless Motion Correction in K-Space of Structural 3D MR-Images of the Brain". In: *ISMRM 2019*. May 2019.

[16] W. Schlegel and J. Bille, eds. *Medizinische Physik 2: Medizinische Strahlenphysik*. Springer-Verlag, 2002. ISBN: 978-3-642-62981-5. DOI: 10.1007/978-3-642-56259-4.

[17] Gunnar Brix et al. *Basics of Magnetic Resonance Imaging and Magnetic Resonance Spectroscopy*. Ed. by Hricak H. Reiser M., Semmler W. Springer, Berlin, Heidelberg, 2008. ISBN: 978-3-540-29355-2. DOI: https://doi.org/10.1007/978-3-540-29355-2_2. URL: https://www.jaypeedigital.com/book/9789351521426/chapter/ch22.

[18] Lars G. Hanson. "Introduction to Magnetic Resonance Imaging Techniques". In: (2009). DOI: http://www.drcmr.dk.

[19] Cynthia B. Paschal and H. Douglas Morris. "K-Space in the Clinic". In: *Journal of Magnetic Resonance Imaging* 19.2 (2004), pp. 145–159. ISSN: 10531807. DOI: 10.1002/jmri.10451.

[20] Maxim Zaitsev, Julian Maclaren, and Michael Herbst. "Motion artifacts in MRI: A complex problem with many partial solutions". In: *Journal of Magnetic Resonance Imaging* 42.4 (2015), pp. 887–901. ISSN: 15222586. DOI: 10.1002/jmri.24850.

[21] Matt A. Bernstein, Kevin F. King, and Xiaohong Joe Zhou. "Chapter 12 - Basics of physiologic gating, triggering, and monitoring". In: *Handbook of MRI Pulse Sequences*. Ed. by Matt A. Bernstein, Kevin F. King, and Xiaohong Joe Zhou. Burlington: Academic Press, 2004, pp. 443 –490. ISBN: 978-0-12-092861-3. DOI: https://doi.org/10.1016/B978-012092861-3/50018-2. URL: http://www.sciencedirect.com/science/article/pii/B9780120928613500182.

[22] Julian Maclaren et al. "Prospective motion correction in brain imaging: A review". In: *Magnetic Resonance in Medicine* 69.3 (2013), pp. 621–636. ISSN: 07403194. DOI: 10.1002/mrm.24314.

[23] Alfred Stadler et al. "Artifacts in body MR imaging: their appearance and how to eliminate them". In: *European Radiology* 17.5 (2007), pp. 1242–1255. ISSN: 1432-1084. DOI: 10.1007/s00330-006-0470-4. URL: https://doi.org/10.1007/s00330-006-0470-4.

[24] Nathan White et al. "PROMO: Real-time prospective motion correction in MRI using image-based tracking". In: *Magnetic Resonance in Medicine* 63.1 (2010), pp. 91–105. ISSN: 07403194. DOI: 10.1002/mrm.22176.

[25] Thomas Benner, André J.W. Van Der Kouwe, and A. Gregory Sorensen. "Diffusion imaging with prospective motion correction and reacquisition". In: *Magnetic Resonance in Medicine* 66.1 (2011), pp. 154–167. ISSN: 15222594. DOI: 10.1002/mrm.22837.

[26] Jure Zbontar et al. "fastMRI: An Open Dataset and Benchmarks for Accelerated MRI". In: (2018), pp. 1–35. arXiv: 1811.08839. URL: http://arxiv.org/abs/1811.08839.

[27] Haimiao Zhang and Bin Dong. "A Review on Deep Learning in Medical Image Reconstruction". In: (2019). arXiv: 1906.10643. URL: http://arxiv.org/abs/1906.10643.

[28] Thomas Küstner et al. "Retrospective correction of motion-affected MR images using deep learning frameworks". In: *Magnetic Resonance in Medicine* 82.4 (2019), pp. 1527–1540. ISSN: 15222594. DOI: 10.1002/mrm.27783.

[29] Alexander Loktyushin et al. "Blind retrospective motion correction of MR images". In: *Magnetic Resonance in Medicine* 70.6 (2013), pp. 1608–1618. ISSN: 07403194. DOI: 10.1002/mrm.24615.

[30] Gitte M. Knudsen et al. "The Center for Integrated Molecular Brain Imaging (Cimbi) database". In: *NeuroImage* 124 (2016). Sharing the wealth: Brain Imaging Repositories in 2015, pp. 1213–1219. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2015.04.025. URL: https://www.sciencedirect.com/science/article/pii/S1053811915003158.

[31] Andreea-Veronica Vascan. "How do children move in the MR scanner with and without anaesthesia?" MA thesis. Department of Computer Science: University of Copenhagen, 2020.

[32] Hannah Susanne Eichhorn et al. "Characterisation of children's head motion for Magnetic Resonance Imaging with and without general anaesthesia". In preparation. In preparation.

[33] *retroMoCoBox, Toolbox for Matlab*. https://github.com/dgallichan/retroMoCoBox. Accessed: 2021-04-18.

[34] John Ashburner. "SPM: A history". In: *NeuroImage* 62.2 (2012). 20 YEARS OF fMRI, pp. 791–800. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2011.10.025. URL: https://www.sciencedirect.com/science/article/pii/S1053811911011888.

[35] Steven Kecskemeti et al. "Robust motion correction strategy for structural MRI in unsedated children demonstrated with three-dimensional radial MPnRAGE". In: *Radiology* 289.2 (2018), pp. 509–516. ISSN: 15271315. DOI: 10.1148/radiol.2018180180.

[36] John P Mugler and James R Brookeman. "Rapid Three-dimensional T1-weighted MR Imaging with MP-RAGE Sequence". In: *Journal of Magnetic Resonance Imaging* 1.5 (1991), pp. 561–567. URL: http://mriquestions.com/uploads/3/4/5/7/34572113/mugler{_}mprage539244.pdf.

[37] Andrea Bernasconi et al. "Recommendations for the use of structural magnetic resonance imaging in the care of patients with epilepsy: A consensus report from the International League Against Epilepsy Neuroimaging Task Force". In: *Epilepsia* 60.6 (2019), pp. 1054–1068. DOI: https://doi.org/10.1111/epi.15612. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/epi.15612. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.15612.

[38] Benjamin M. Ellingson et al. "Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials". In: *Neuro-Oncology* 17.9 (2015), pp. 1188–1198. ISSN: 15235866. DOI: 10.1093/neuonc/nov095.

[39] John P. Mugler. "Optimized three-dimensional fast-spin-echo MRI". In: *Journal of Magnetic Resonance Imaging* 39.4 (2014), pp. 745–767. ISSN: 15222586. DOI: 10.1002/jmri.24542.

[40] Matthew T. Walker et al. "Fast, versatile, and cost-effective FSE MR imaging: technical considerations and clinical applications". In: *Barrow Quarterly* 16.2 (2000), pp. 1–5.

[41] Glenn Krinsky, Neil M. Rofsky, and Jeffrey C. Weinreb. "Nonspecificity of short inversion time inversion recovery (STIR) as a technique of fat suppression: Pitfalls in image interpretation". In: *American Journal of Roentgenology* 166.3 (1996), pp. 523–526. ISSN: 0361803X. DOI: 10.2214/ajr.166.3.8623620.

[42] G.M. Bydder and I.R. Young. "Clinical use of the partial saturation and saturation recovery sequences in MR imaging". In: *Journal of Computer Assisted Tomography* 9.4 (1985), pp. 659–675. DOI: 10.1097/00004728-198511000-00004.

[43] R E Osborn et al. "MR imaging of neuronal migrational disorders." In: *American Journal of Neuroradiology* 9.6 (1988), pp. 1101–1106. ISSN: 0195-6108. eprint: http://www.ajnr.org/content/9/6/1101.full.pdf. URL: http://www.ajnr.org/content/9/6/1101.

[44] David J. Mikulis and Timothy P.L. Roberts. "Neuro MR: Protocols". In: *Journal of Magnetic Resonance Imaging* 26.4 (2007), pp. 838–847. ISSN: 10531807. DOI: 10.1002/jmri.21041.

[45] O. Naggara et al. "Diffusion-weighted MR imaging of the brain". In: *Journal de Radiologie* 91.3 C2 (2010), pp. 329–351. ISSN: 02210363. DOI: 10.1016/S0221-0363(10)70050-0.

[46] Jakob M. Slipsager et al. "Markerless motion tracking and correction for PET, MRI, and simultaneous PET/MRI". In: *PLoS ONE* 14.4 (2019). ISSN: 19326203. DOI: 10.1371/journal.pone.0215524.

[47] Oline V. Olesen et al. "Structured Light 3D Tracking System for Measuring Motions in PET Brain Imaging". In: *Proceedings of SPIE Medial Imaging - The International Society for Optical Engineering*. Vol. 7625. 2010, pp. 121–135.

[48] Bruce Fischl. "FreeSurfer". In: *NeuroImage* 62.2 (2012), pp. 774–781. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2012.01.021.

[49] Douglas N. Greve and Bruce Fischl. "Accurate and robust brain image alignment using boundary-based registration". In: *NeuroImage* 48.1 (2009), pp. 63–72. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2009.06.060. URL: http://dx.doi.org/10.1016/j.neuroimage.2009.06.060.

[50] Martin Reuter, H. Diana Rosas, and Bruce Fischl. "Highly accurate inverse consistent registration: A robust approach". In: *NeuroImage* 53.4 (2010), pp. 1181–1196. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.07.020. URL: http://dx.doi.org/10.1016/j.neuroimage.2010.07.020.

[51] Zhou Wang et al. "Image quality assessment: From error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. ISSN: 10577149. DOI: 10.1109/TIP.2003.819861.

[52] Eric Krotkov. "Focusing". In: *International Journal of Computer Vision* 1.3 (1988), pp. 223–237. ISSN: 09205691. DOI: 10.1007/BF00127822.

[53] Said Pertuz, Domenec Puig, and Miguel Angel Garcia. "Analysis of focus measure operators for shape-from-focus". In: *Pattern Recognition* 46.5 (2013), pp. 1415–1432. ISSN: 00313203. DOI: 10.1016/j.patcog.2012.11.011. URL: http://dx.doi.org/10.1016/j.patcog.2012.11.011.

[54] Johan Berglund et al. "Prospective motion correction for diffusion weighted EPI of the brain using an optical markerless tracker". In: *Magnetic Resonance in Medicine* 85.3 (2020), pp. 1427–1440. ISSN: 15222594. DOI: 10.1002/mrm.28524.

[55] Gregory W. Corder and Dale I. Foreman. *Nonparametric Statistics: a Step-by-Step Approach*. 2nd Editio. John Wiley & Sons, Inc., 2014. ISBN: 978-1-118-84031-3.

[56] Frank Wilcoxon. "Individual Comparisons by Ranking Methods". In: *Biometrics Bulletin* 1.6 (1945), pp. 80–83. DOI: 10.1007/978-1-4612-4380-9_16.

[57] *Scipy Documentation*. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.wilcoxon.html. Accessed: 2021-05-4.

[58] Yoav Benjamini and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 57.1 (1995), pp. 289–300. DOI: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x. eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1995.tb02031.x. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1995.tb02031.x.

[59] Steven A. Julious. "Tutorial in biostatistics: Sample sizes for clinical trials with Normal data". In: *Statistics in Medicine* 23.12 (2004), pp. 1921–1986. ISSN: 02776715. DOI: 10.1002/sim.1783.

[60] Nathan W. Churchill et al. "Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. Standard temporal motion and physiological noise correction methods". In: *Human Brain Mapping* 33.3 (2012), pp. 609–627. ISSN: 10659471. DOI: 10.1002/hbm.21238.

[61] Sune H. Keller et al. "Motion correction in simultaneous PET/MR brain imaging using sparsely sampled MR navigators: a clinically feasible tool". In: *EJNMMI Physics* 2.1 (2015), pp. 1–7. ISSN: 21977364. DOI: 10.1186/s40658-015-0118-z. URL: http://dx.doi.org/10.1186/s40658-015-0118-z.

[62] Kiaran P. McGee et al. "Image metric-based correction (Autocorrection) of motion effects: Analysis of image metrics". In: *Journal of Magnetic Resonance Imaging* 11.2 (2000), pp. 174–181. ISSN: 10531807. DOI: 10.1002/(SICI)1522-2586(200002)11:2<174::AID-JMRI15>3.0.CO;2-3.

[63] Kaiyang Cheng et al. "Addressing The False Negative Problem of Deep Learning MRI Reconstruction Models by Adversarial Attacks and Robust Training". In: *Proceedings of Machine Learning Research*. Vol. 121. 2020, pp. 121–135. URL: https://slideslive.com/38922093/medical-imaging-meets-neurips-4.

[64] Muhammad Usman et al. "Retrospective Motion Correction in Multishot MRI using Generative Adversarial Network". In: *Scientific Reports* 10.1 (2020), pp. 1–11. ISSN: 20452322. DOI: 10.1038/s41598-020-61705-9. URL: http://dx.doi.org/10.1038/s41598-020-61705-9.

[65] Alain Horé and Djemel Ziou. "Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?" In: *IET Image Processing* 7.1 (2013), pp. 12–24. ISSN: 17519659. DOI: 10.1049/iet-ipr.2012.0489.

[66] Jakob M Slipsager et al. "Comparison of Prospective and Retrospective Motion Correction for 3D Structural Brain MRI". In: *ISMRM 2020*. 2020.

[67] Benjamin M. Kozak et al. "Mri techniques to decrease imaging times in children". In: *Radiographics* 40.2 (2020), pp. 485–502. ISSN: 15271323. DOI: 10.1148/rg.2020190112.

[68] James G. Pipe. "Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging". In: *Magnetic Resonance in Medicine* 42.5 (1999), pp. 963–969. ISSN: 07403194. DOI: 10.1002/(SICI)1522-2594(199911)42:5<963::AID-MRM17>3.0.CO;2-L.

[69] Leo F. Czervionke et al. "Characteristic Features of MR Truncation Artifacts". In: *American Journal of Neuroradiology* 9.5 (1988), pp. 815–824. ISSN: 0195-6108. eprint: http://www.ajnr.org/content/9/5/815.full.pdf. URL: http://www.ajnr.org/content/9/5/815.

[70] Martin Reuter et al. "Within-subject template estimation for unbiased longitudinal image analysis". In: *NeuroImage* 61.4 (2012), pp. 1402–1418. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2012.02.084. URL: https://www.sciencedirect.com/science/article/pii/S1053811912002765.

[71] Vegard Antun et al. "On instabilities of deep learning in image reconstruction and the potential costs of AI". In: *Proceedings of the National Academy of Sciences* (2020), p. 201907377. ISSN: 0027-8424. DOI: 10.1073/pnas.1907377117.

A Additional material for Methods

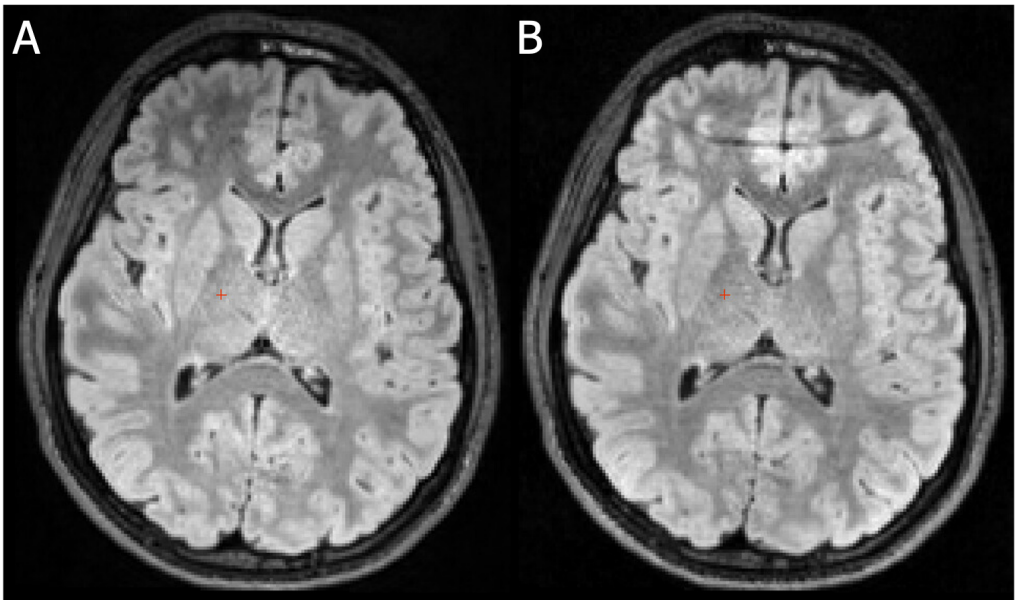


FIGURE A.1: Example of artefact introduced by the retrospective reconstruction algorithm in FLAIR scans. A scan for nodding motion with pMC and reacquisition is shown, which is reconstructed (A) by the scanner and (B) retrospectively.

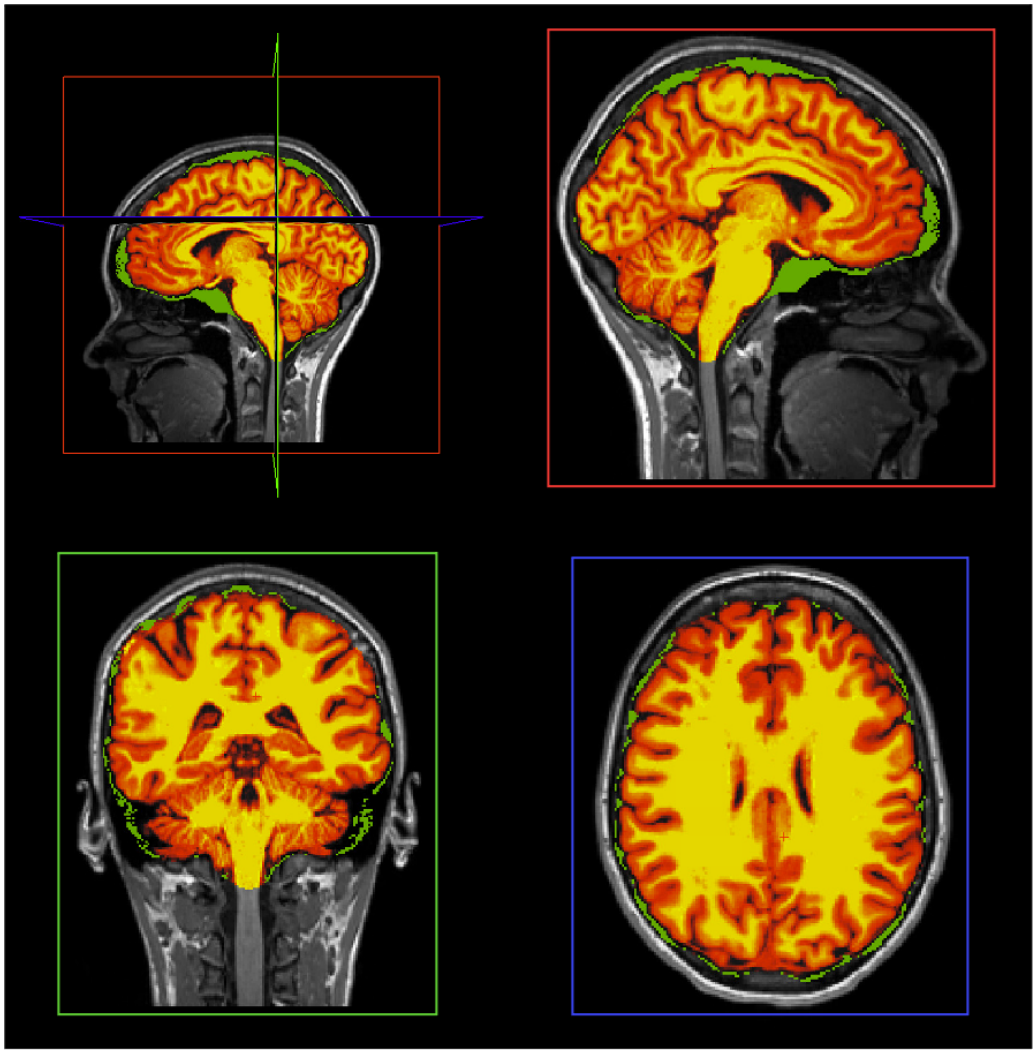


FIGURE A.2: Example of brain mask edit using the graph cut option in FreeSurfer's *recon-all*. In the visualisation, the brainmask (yellow) is overlayed with the MPRAGE scan, green indicates the voxels which are removed from the brainmask after editing.

B Additional material for Results

B.1 Motion patterns

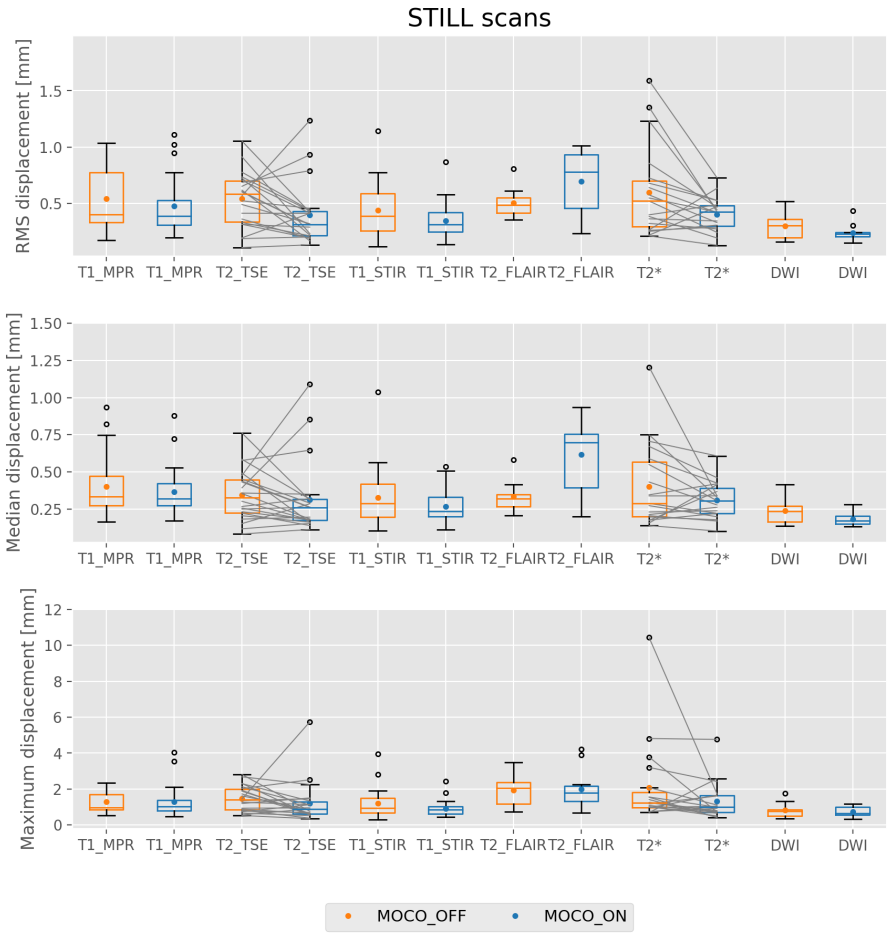


FIGURE B.1: Comparison of the RMS, median and maximum displacement with respect to the mid of acquisition for the motionless scans. Median values of the respective metrics are visualised as box plots, mean values are shown as orange and blue dots. No statistical significant differences were observed after FDR correction at `FDR=0.05`.

B.2 Image quality measures

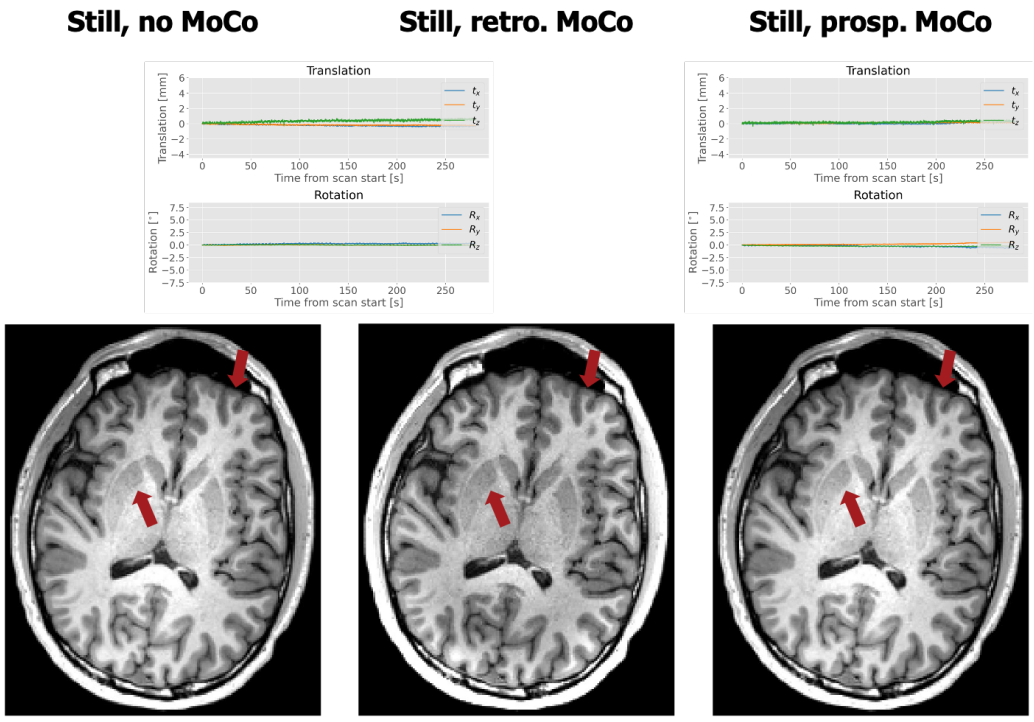


FIGURE B.2: Exemplary T_1 MPRAGE images, comparing the still scans for all motion correction settings. The differences between the individual images are too small to see with the human eye. The corresponding motion curves are shown above the images. Since rMC was applied to the uncorrected scans, the motion curve in the middle corresponds to both images. All images shown are registered to the still, uncorrected scan.

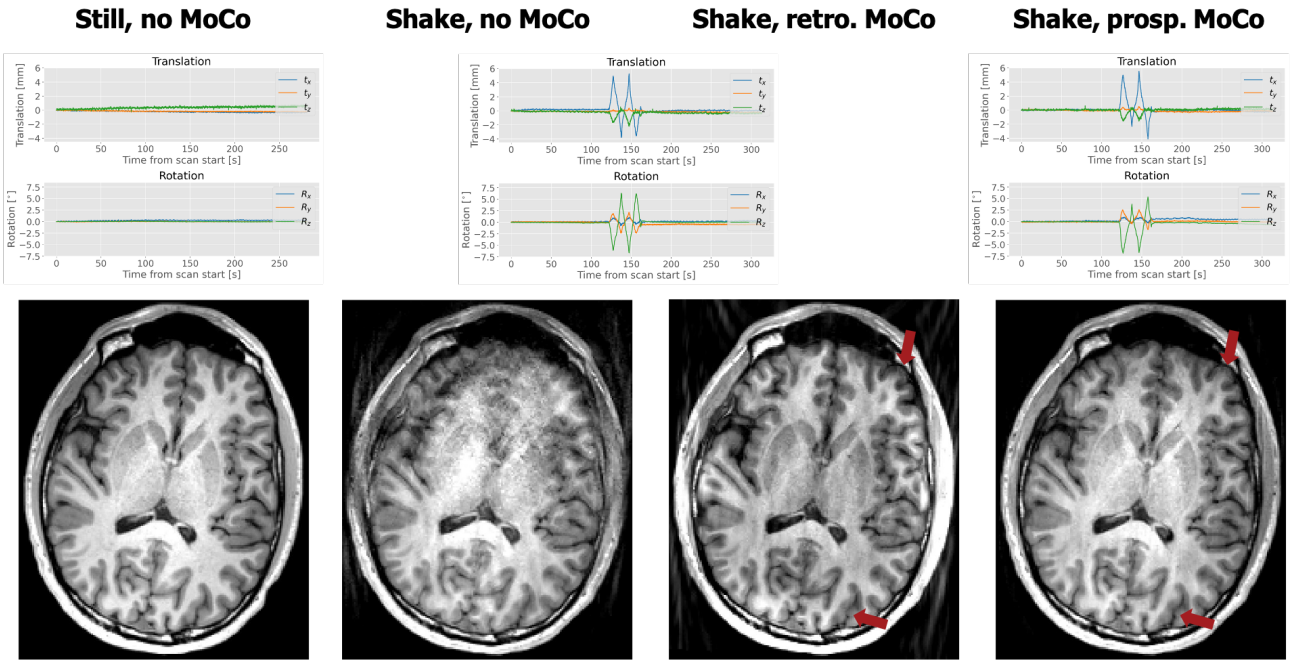


FIGURE B.3: Exemplary T_1 MPRAGE images, comparing the shaking scans for all motion correction settings with the ground truth scan (left). The corresponding motion curves are shown above the images. Since rMC was applied to the uncorrected scans, the motion curve in the middle corresponds to both images. All images shown are registered to the still, uncorrected scan.

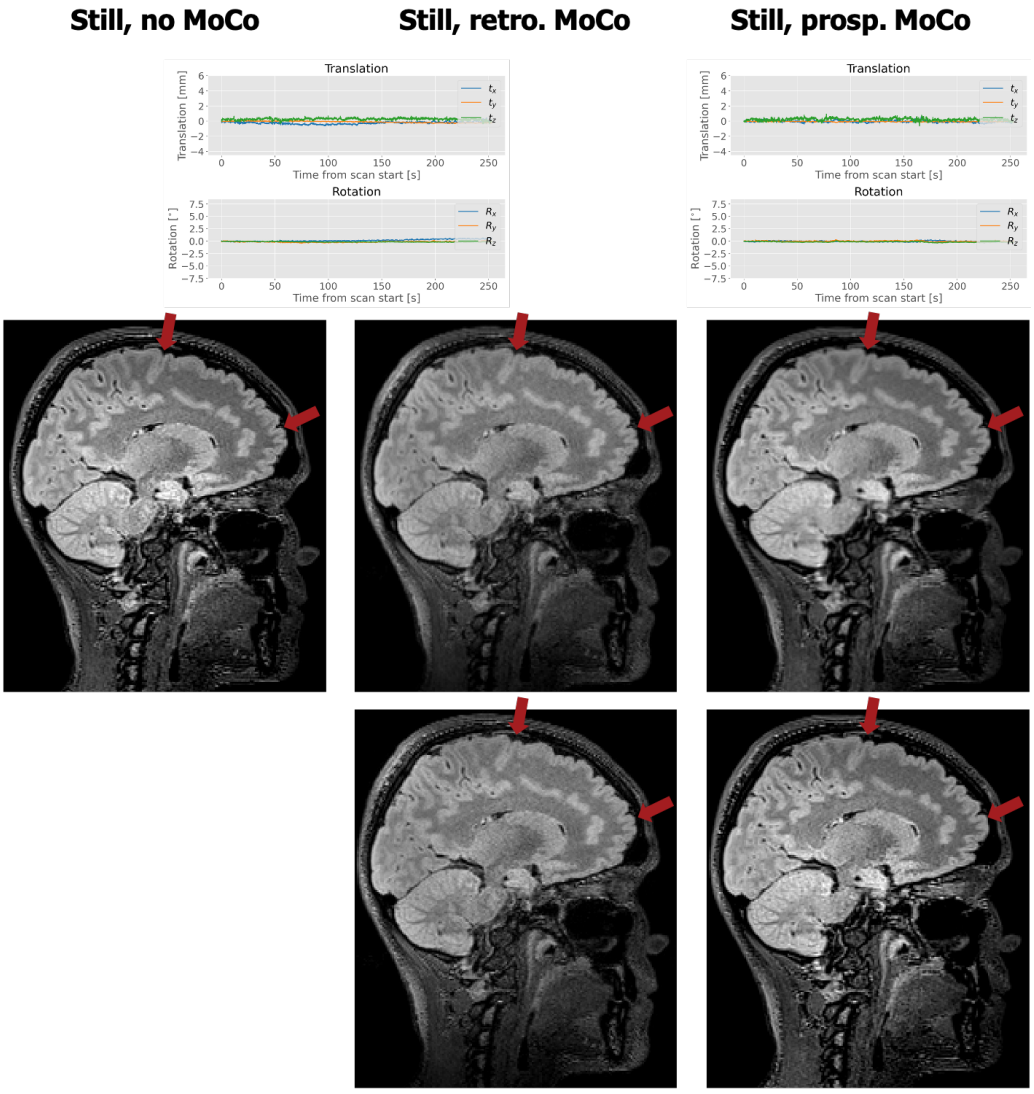


FIGURE B.4: Exemplary T_2 FLAIR images, comparing the still scans for all motion correction settings. Images in the first row are registered to the still, uncorrected scan. Images in the second row are not registered, in order to examine the effects without interpolation related smoothing. The differences between the individual images are quite small and difficult to appreciate in individual slices next to each other. However, red arrows indicate regions, where white and grey matter boundaries for the prospectively corrected scan are slightly more blurry than for the other two scans. This effect is reduced, when comparing unregistered images. However, not all blurring can be explained by interpolation. The corresponding motion curves are shown above the images. Since rMC was applied to the uncorrected scans, the motion curve in the middle corresponds to both images.

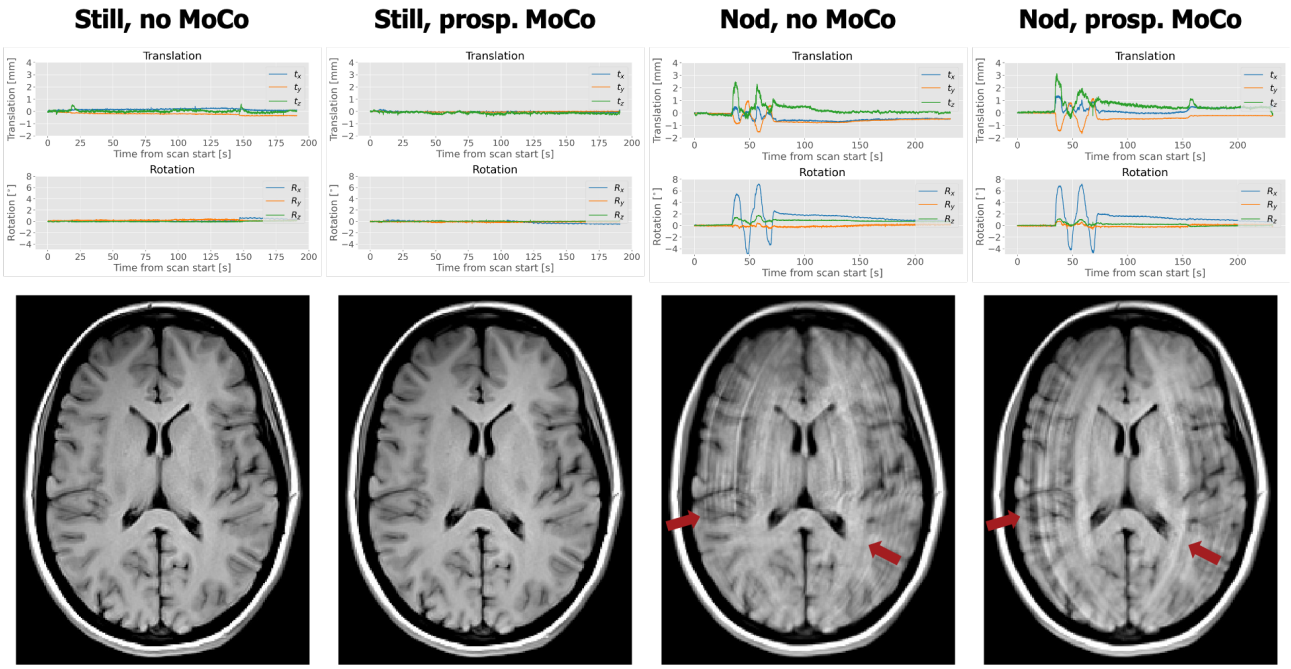


FIGURE B.5: Example of pMC not improving image quality in T_1 STIR images. The still and nodding scans with and without pMC are compared (without reacquisition). Red arrows indicate areas, where pMC introduces severe additional ringing in the nodding scan. All images shown are registered to the still, uncorrected scan.

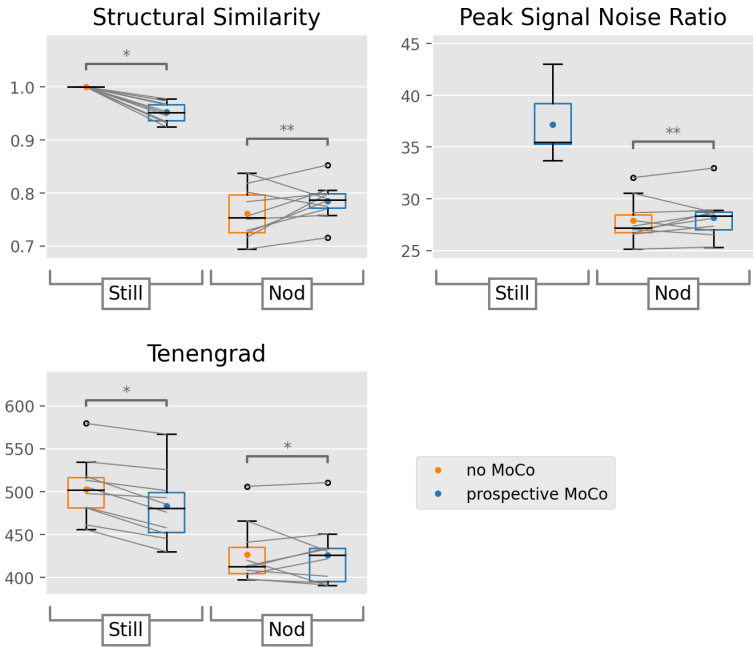


FIGURE B.6: Comparison of image quality metrics SSIM, PSNR and Tenengrad for prospectively corrected and uncorrected DWI TRACE weighted scans with $b = 0\frac{s}{mm^2}$ (without reacquisition). Statistical significance after FDR correction (of all comparisons) is indicated by */** ($p < 0.05/p < 0.001$). The individual paired data points are connected with grey lines. All image quality measures show a significant, but small increase of image quality for the motion scans with pMC compared to the uncorrected scans.

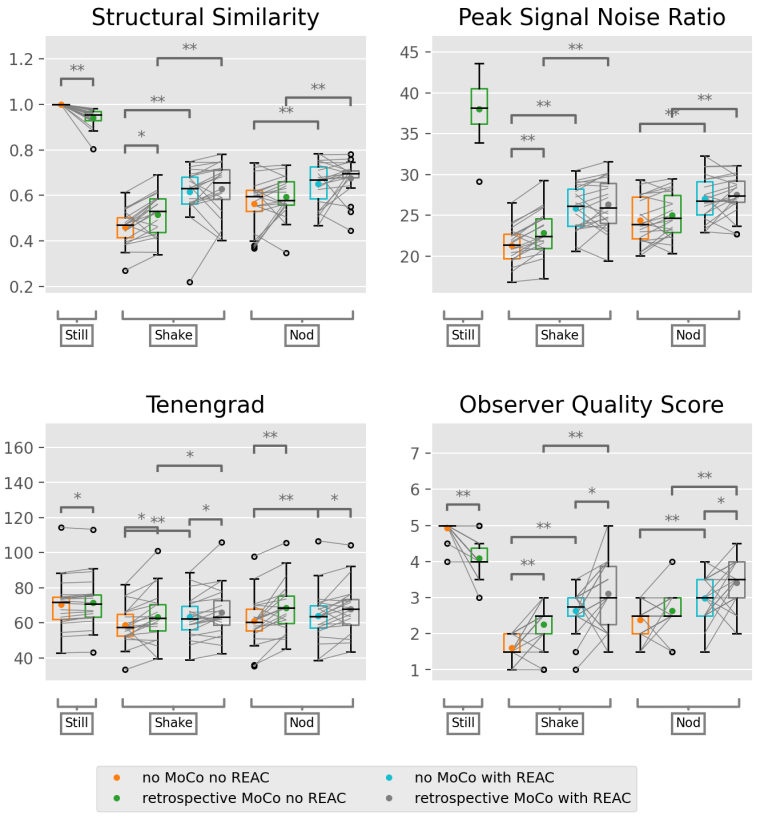


FIGURE B.7: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_1 MPRAGE scans with and without rMC as well as reacquisition, for all types of motion patterns. Statistical significance after FDR correction (of all comparisons) is indicated by */** ($p < 0.05/p < 0.001$). The individual paired data points are connected with grey lines.

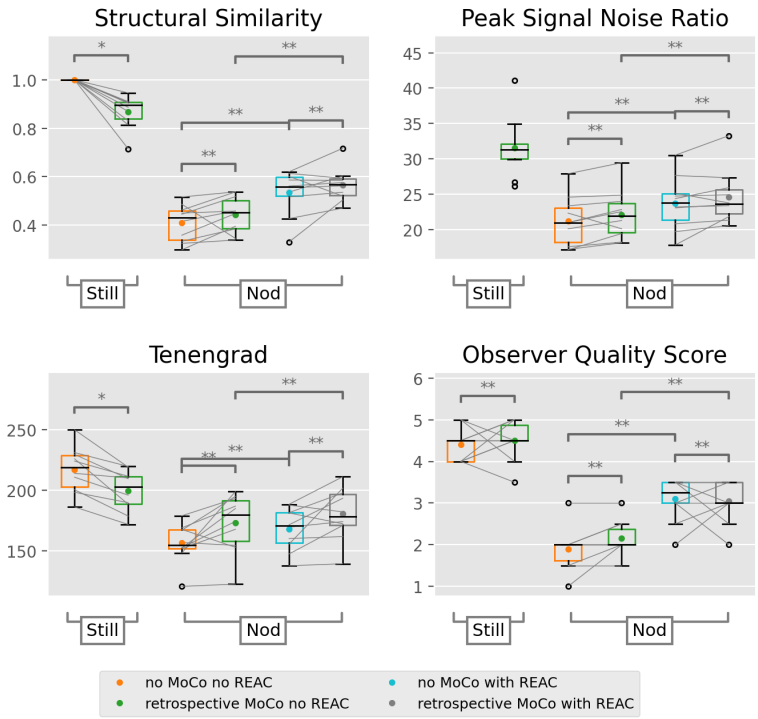


FIGURE B.8: Comparison of image quality metrics SSIM, PSNR and Tenengrad, as well as observer quality scores for T_2 FLAIR scans with and without rMC as well as reacquisition, for all types of motion patterns. Statistical significance after FDR correction (of all comparisons) is indicated by */** ($p < 0.05/p < 0.001$). The individual paired data points are connected with grey lines.

B.3 Cortical thickness changes

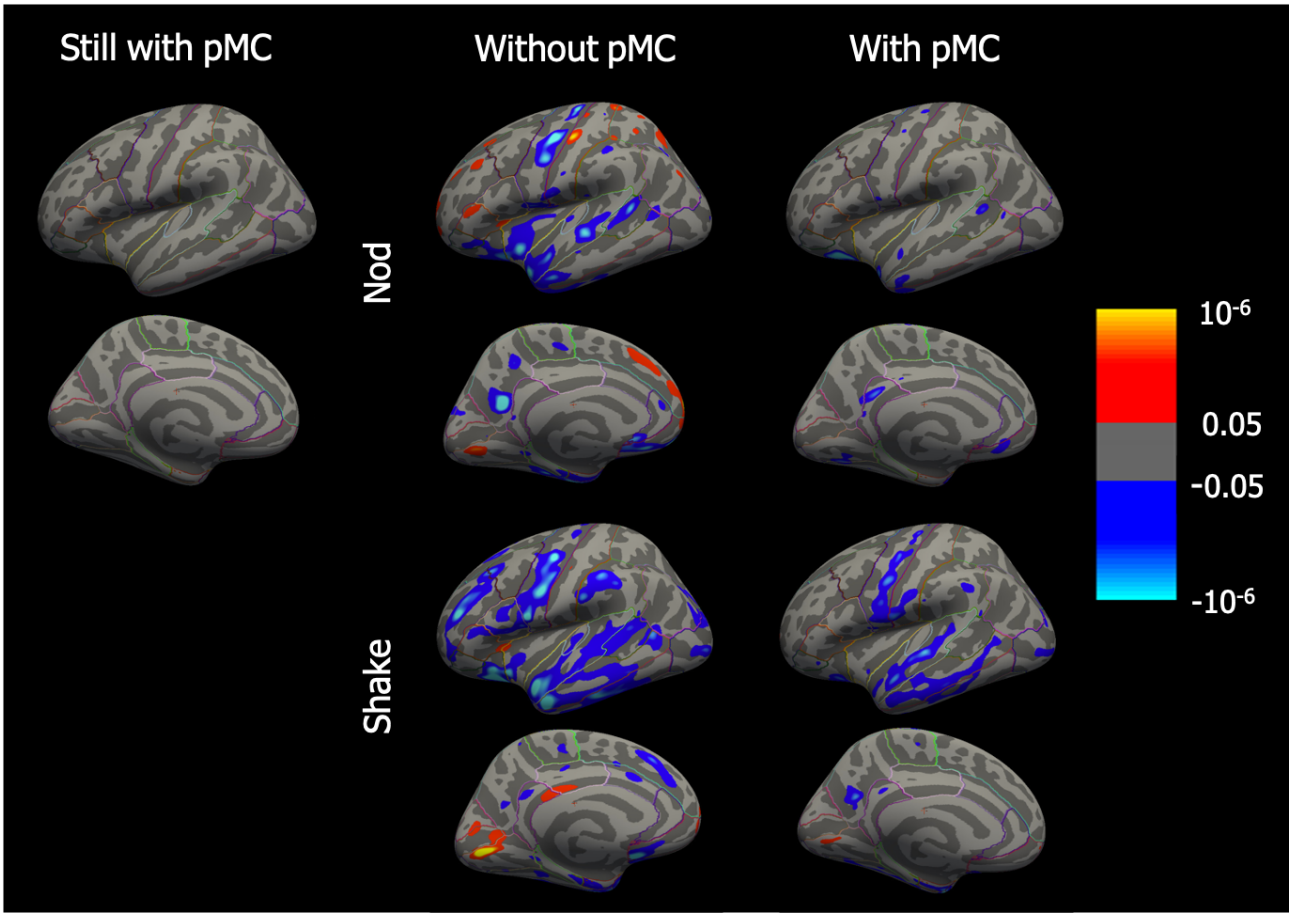


FIGURE B.9: Significance maps of the left hemisphere for GLM analysis of cortical thickness changes with increased RMS displacement in nodding and motionless scans processed with *FreeSurfer's* longitudinal processing stream. Blue represents decreasing cortical thickness, yellow and red increasing thickness with increasing motion. The visualised p-values were FDR thresholded at `FDR=0.05` for both hemispheres. pMC does not introduce a bias on motionless scans. However, for both nodding and shaking motion pMC decreases the number of significant vertices. Scans with nodding motion do not show as many significant changes as scans with shaking motion. Compared to the analogous analysis of scans processed with the cross-sectional stream (Figures 41 and B.11), the amount of significant vertices is decreased by first estimating surfaces on a within-subject template and then fine adjusting the estimates for each individual scan in the longitudinal processing.

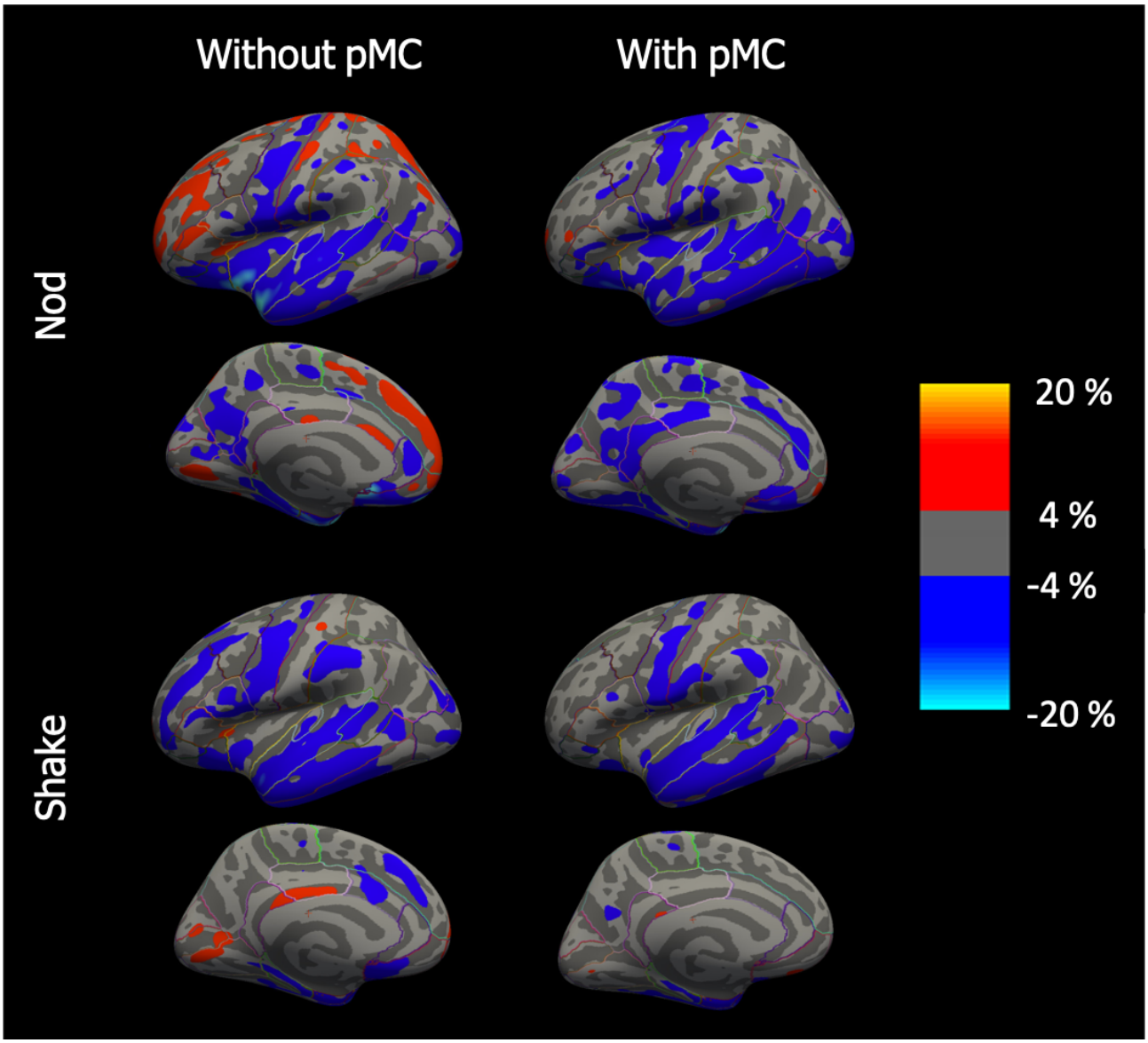


FIGURE B.10: Percentage thickness changes for increased RMS displacements for scans showing a correlation in Figure B.9 (processed with *FreeSurfer's* longitudinal stream). Blue represents a negative slope, yellow and red a positive one. The average cortical thickness over all cortical areas and all subjects processed in the longitudinal stream is 2.43 ± 0.07 mm. Thus, a change of 4 % (dark blue or red) per 1 mm increased RMS displacement corresponds to approximately 0.10 mm thickness change, whereas a change of 20 % (light blue or yellow) corresponds to approximately 0.49 mm. The majority of vertices for scans without pMC and all visualised vertices for scans with pMC only show minor thickness changes. The number of vertices with changes above 4 % is decreased by pMC. In comparison with results from the cross-sectional processing stream (Figures 42 and B.12), processing with the longitudinal stream lead to a reduced number of vertices with changes above 4 %.

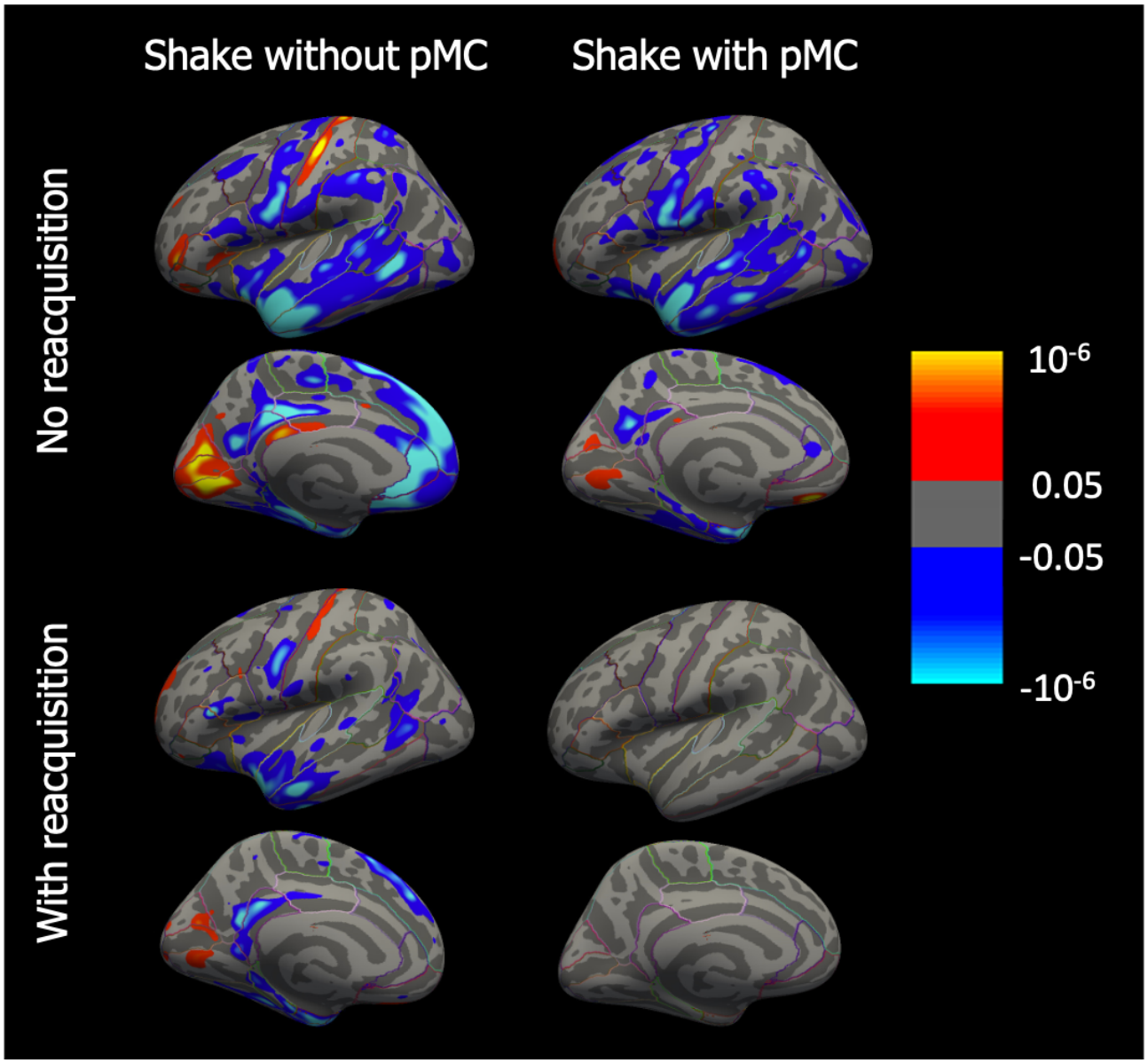


FIGURE B.11: Significance maps of the left hemisphere for GLM analysis of cortical thickness changes with increased RMS displacement in scans with shaking motion. Blue represents decreasing cortical thickness, yellow and red increasing thickness with increasing motion. The visualised p-values were FDR thresholded at `FDR=0.05` for both hemispheres. pMC and reacquisition both decrease the number of significant vertices. However, pMC is not able to reduce the amount of significant thickness changes to the same extent as for nodding scans (shown in Figure 41). The combination of pMC and reacquisition again does not show any statistically significant correlation with motion.

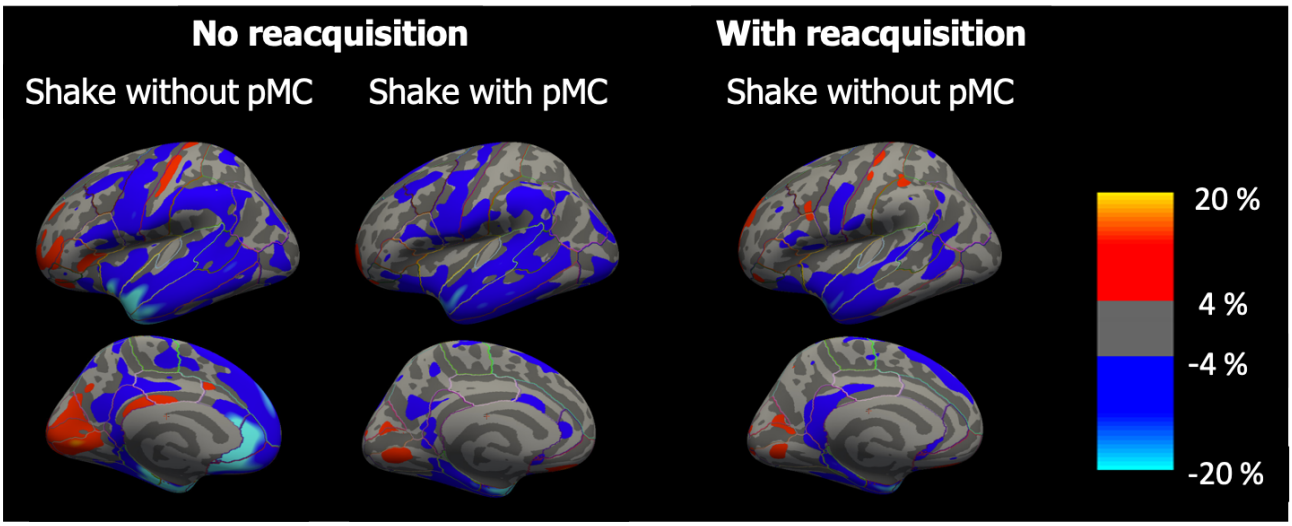
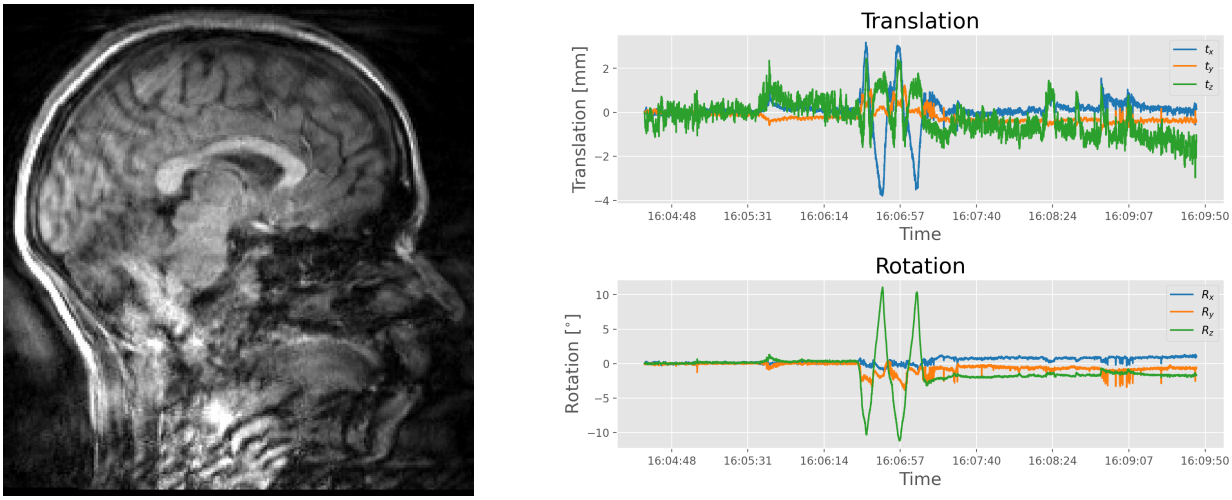


FIGURE B.12: Percentage thickness changes for increased RMS displacements for scans showing a correlation in Figure B.11. Blue represents a negative slope, yellow and red a positive one. The average cortical thickness over all subjects and all cortical areas is 2.40 ± 0.07 mm. Thus, a change of 4 % (dark blue or red) per 1 mm increased RMS displacement corresponds to approximately 0.10 mm thickness change, whereas a change of 20 % (light blue or yellow) corresponds to approximately 0.48 mm. The larger part of vertices for all scan types only shows very small thickness changes. The number of vertices with changes above 4 % is decreased by reacquisition, as well as pMC. The amount of vertices with large slopes is comparable with the results for nodding motion in Figure 42.

C Additional material for Discussion



(A) (B)

FIGURE C.1: Outliers with severe ringing in some prospectively corrected shaking scans without reacquisition: (A)an exemplary image in sagittal view and (B) the corresponding motion curves showing peaks, where the participant was lying still, but deforming his/her face non-rigidly.

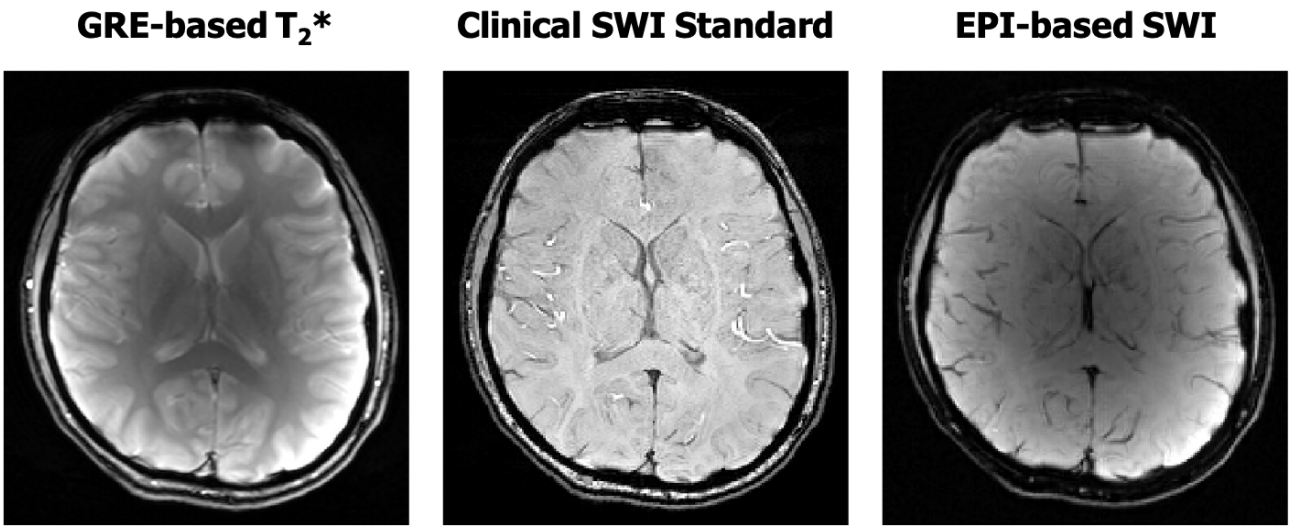


FIGURE C.2: Comparison of GRE-based T_2* sequence with the clinical standard GRE-based SWI sequence, as well as the accelerated EPI-based SWI sequence. The images have similar information content, especially about vessels.

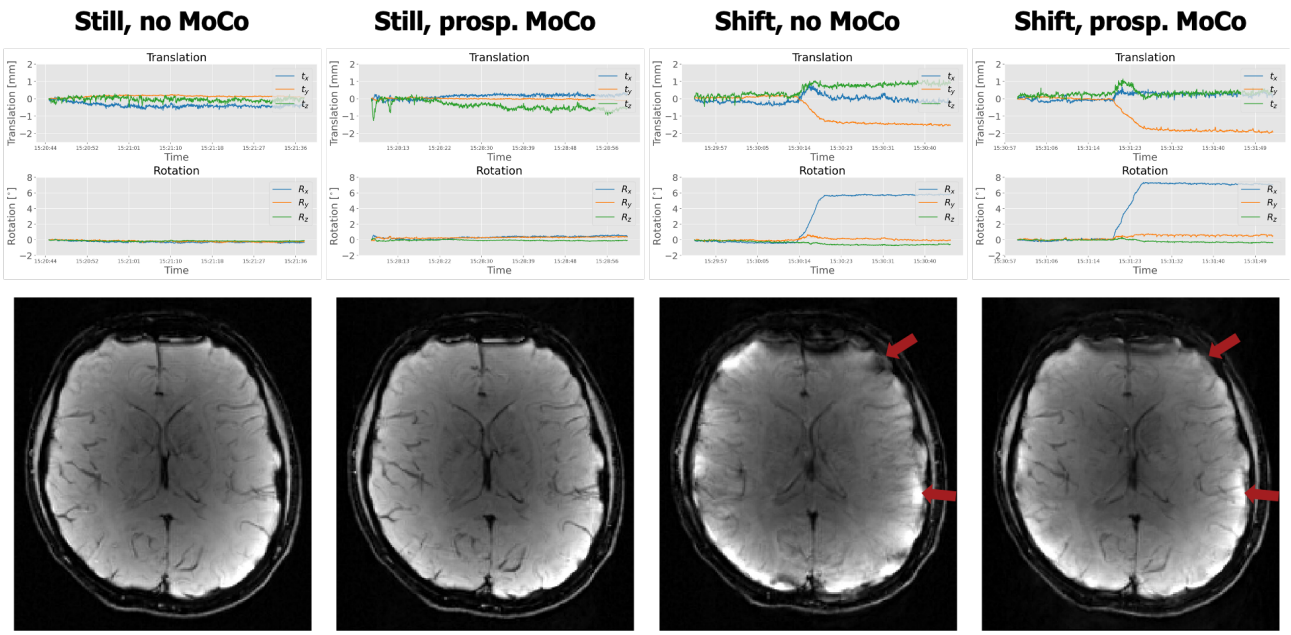


FIGURE C.3: Example of pMC for EPI-based SWI sequence for motionless scans as well as scans with a shift or rather, rotation out of the original position. Red arrows indicate, where pMC reduced blurring and signal loss artefacts. The corresponding motion curves are shown above the images.

D Preprint of article "Characterisation of children's head motion for Magnetic Resonance Imaging with and without general anaesthesia" (in preparation)

Characterisation of children's head motion for Magnetic Resonance Imaging with and without general anaesthesia

Hannah Susanne Eichhorn[a,b], Andreea-Veronica Vascan[a,c], Martin Nørgaard[a,d], Andreas Høyby Ellegaard[e], Jakob Slipsager[f,g,e], Sune Høgild Keller[g], Lisbeth Marner[g,h], Melanie Ganz[a,c,*]

[a] *Neurobiology Research Unit, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark*
[b] *Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark*
[c] *Department of Computer Science, University of Copenhagen, Copenhagen, Denmark*
[d] *Center for Reproducible Neuroscience, Department of Psychology, Stanford University, Stanford, CA, USA*
[e] *TracInnovations, Ballerup, Denmark*
[f] *DTU Compute, Technical University of Denmark, Lyngby, Denmark*
[g] *Department of Clinical Physiology, Nuclear Medicine & PET, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark*
[h] *University Hospital Bispebjerg, Copenhagen, Denmark.*

Abstract

Head motion is one of the major reasons for artefacts in Magnetic Resonance Imaging (MRI), which is especially challenging for children who are often intimidated by the dimensions of the MR scanner. In order to optimise the MRI acquisition for children in the clinical setting, insights into children's motion patterns are essential. In this work, we analyse motion data from 61 pediatric patients. We compare structural MRI data of children imaged with and without general anaesthesia (GA), all scanned using the same hybrid PET/MR scanner. We analyse several metrics of motion based on the displacement relative to a reference, decompose the transformation matrix into translation and rotation, as well as investigate how different regions in the brain are affected by motion. Head motion for children without GA was significantly higher (mean displacement of 2.19 ± 0.93 mm (median $\pm$ standard deviation) during 41.7 ± 7.5 min

*Corresponding author
Email address: melanie.ganz@nru.dk (Melanie Ganz)

scans); however, even anaesthetised children showed substantial residual head motion (mean displacement of 1.12±0.35 mm). For both patient groups translation along the z-axis (along the scanner bore) was significantly larger in absolute terms (GA / no GA: 0.87 ± 0.29 mm / 0.92 ± 0.49 mm) compared to the other directions. Considering directionality, both patient groups were moving in negative z-direction and thus, out of the scanner. The awake children additionally showed significantly more nodding rotation (0.33 ± 0.20 °). Consequently, in future studies as well as in the clinical setting, these predominant types of motion need to be taken into consideration to limit artefacts and reduce re-scans due to poor image quality.

Keywords: Magnetic Resonance Imaging, Motion patterns, Motion Artefact, Children, Brain

1. Introduction

For Magnetic Resonance Imaging (MRI), artefacts are most frequently caused by patient head motion due to long acquisition times of typically 30 to 60 minutes [1]. These artefacts manifest as ghosting, blurring or signal variations, thus reducing overall image quality and resulting in unsuccessful diagnoses [2]. Andre et.al. [3] determined the percentage of at least partly repeated MRI examinations to approximately 20%, leading to an estimated increased cost of 115,000 US dollars per scanner per year in the US due to motion.

Children tend to move more in the MR scanner than adults since they are more negatively influenced by the large size of the scanner, the narrow bore and the loud noises during image acquisition [4]. So far, motion artefacts for children between 4 and 10 years are mostly reduced by sedation or general anaesthesia (GA) [5], with the exact age limit being hospital specific. However, the need for an anaesthetist to administer the drug and monitor the child, increases the patient's waiting times, as well as the costs of the examination. Slipsager et.al. [6] calculated the additional cost for using GA in MRI examinations to 319,000 US dollars per scanner per year in Denmark. Furthermore, concerns about

adverse events like airway obstruction or oxygen desaturation of GA in young children are increasing [1, 4, 7, 8].

Currently, different approaches for preventing motion artefacts without general anaesthesia or sedation are under research. These include strategies for preparing the children with story books and mock scanners before they undergo MRI and distracting them visually and acoustically during the examination [4, 5, 9, 10]. Another approach is to apply motion correction during or after image acquisition, for which a variety of methods have been developed and tested. These include prospective motion correction updating the position of the field of view in real time dependent on motion estimates [11, 12, 13], as well as several retrospective techniques using mathematical properties of the Fourier Transform, Compressed Sensing or Machine Learning [14, 15, 16, 17].

For both strategies - preparation and motion correction - information about the children's motion patterns in the MR scanner is essential. In the case of motion correction algorithms, additional information is needed for tailoring them towards children specific movements. For the preparation strategy as well as in the clinical setting, information about predominant motion patterns is critical for preventing those types of motion. Churchill et.al. [18] analysed head motion of adults during fMRI using retrospective image-based motion correction. Overall, they found low motion estimates with only two cases of estimates above 1mm or 1° for the whole brain. The largest standard deviation of the displacement relative to a reference image was observed for pitch movement (nodding). Afacan et.al. [1] investigated children's head motion and its impact on image quality. Their analysis showed a correlation of mean displacement and motion free time to image quality, but no statistically significant correlation of maximum displacement. Additionally, they did not find a significant correlation between the analysed motion metrics and age. Together, these results suggest that head motion is a challenge, not only for young and uncooperative children.

The aim of this work is to analyse motion patterns of children in order to draw possible conclusions about predominant movement habits and thus, enable optimisation of motion correction methods for specific types of motion. For that,

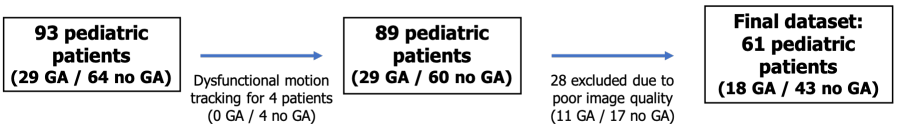


Figure 1: Schematic overview of exclusion process.

Table 1: Demographic information for patient population

	GA	No GA	All
Number of patients (sex)	18 (10 male)	43 (28 male)	61 (38 male)
Mean age ± std	5.94 ± 2.75	12.65 ± 3.04	10.67 ±4.25
Age range	21 m - 13 y	7 y - 19 y	21 m - 19 y

we also investigate how head motion translates into motion of different parts of the brain. The analyses will be performed both for children imaged with and without GA.

2. Material and Methods

2.1. Patient population

Within this study, MR and motion data from 93 pediatric patients with brain tumors acquired in a previous study were analysed [19]. As visualised in Figure 1, four of those data sets were excluded due to dysfunctional motion tracking. Twenty eight additional data sets were excluded due to poor scan quality (ringing or blurring), large tumors or removal of large parts of the brain. Examples of excluded scans as well as a comparison of motion between excluded and included patients are provided in the Supplementary Material S.1 and S.2. Consequently, the final dataset consists of 61 patients aged between 21 months and 19 years, out of which 18 were scanned with GA (propofol and sevoflurane) and 43 were scanned without. Additional demographic information is provided in Table 1. Please note the larger age range of anaesthetised children, which is due to clinical considerations.

Table 2: MR Sequences

Sequence	**Scan Duration [s]**
T_1-weighted MPRAGE	266
T_1-weighted STIR	174
T_2-weighted FLAIR (transversal)	272
T_2-weighted FLAIR (coronal)	164
T_2-weighted Blade	122

The previous study, in which our data was aquired, was approved by the Danish National Committee on Health Research Ethics (approval H-6-2014-095) and was registered at `clinicaltrials.gov` (NCT03402425). Written informed consent was obtained from all patients / parents of the patients.

2.2. MRI acquisition and region segmentation

MR scans were performed using the mMR Biograph hybrid PET/MRI scanner (*Siemens Healthineers, Erlangen, Germany*) between April 2015 and January 2019, using a PET/MR child brain tumor protocol. The median wall time from the first to the last MR sequence was 41.7 ± 7.5 minutes. The MRI sequences of this protocol are listed in Table 2. Parameters of these sequences are reported in the Supplementary Material (Table S.1). The quality of the majority of the 61 scans in the final dataset was assessed as "optimal for clinical use" by radiologists, with only 6.3% of the scans scored "useful for diagnosis, but not optimal", as previously described by Slipsager et.al. [6].

The 3D-encoded, T_1-weighted MPRAGE scans were processed with FreeSurfer [20] in order to segment 8 cortical and 8 subcortical regions: left and right hippocampus, caudate, amygdala, putamen, lateral occipital, inferior temporal, precentral and medial orbitofrontal regions. The centroids of these regions are visualised in Figure 2.

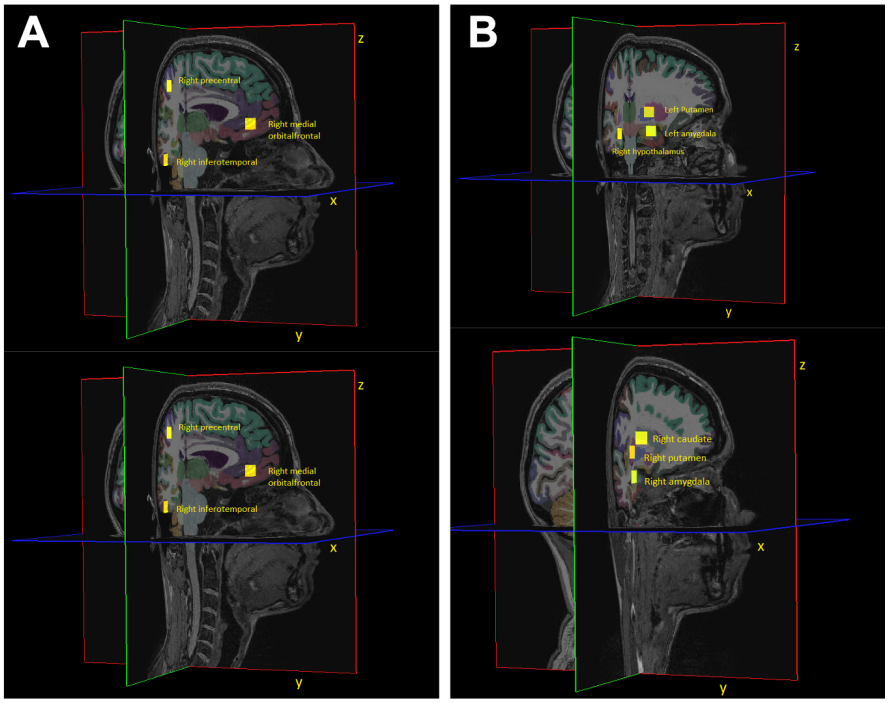


Figure 2: Visualisation of (A) 8 cortical and (B) 8 subcortical regions together with their centroid (yellow square).

2.3. Motion tracking

The patient's head motion was estimated with the markerless tracking system Tracoline (*TracInnovations, Ballerup, Denmark*) [21, 22]. It transmits non-visible infrared light onto the patient's face and estimates a 3D point cloud surface, the position of which is continuously measured at a frequency of approximately 30 Hz. An example of this point cloud for an adult as well as a schematic visualisation of the point cloud centroid is shown in Figure S.3 of the Supplementary Material. Our analysis in the following was only performed on the motion data sampled during acquisition of the sequences summarized in Table 2. For each motion estimate, the middle time point of the MPRAGE acquisition was chosen as the reference position, since this time point corresponds to the acquisition of the center of k-space. Using a calibration transform, the estimates were moved into the RAS (Right, Anterior, Superior) coordinate system.

2.4. Motion quantification

Motion Metrics. In line with the analysis of Afacan et al. [1], the displacement of the point cloud centroid to the reference position was calculated for all time points as Euclidean distance to the reference position. For each patient, the mean, median and maximum of the displacement was determined. Additionally, motion-free time was quantified as the percentage of time, where the displacement relative to the reference position was below 2 mm. This threshold was chosen different from Afacan et al. [1], since 2 mm is a standard value for motion thresholding at our institute.

Analogously, for each region the displacement of the region's centroid (cf. Figure 2) and the corresponding metrics were computed.

Matrix decomposition. Tracoline provides an estimate of the rigid body transformation matrix at each time point. Analogous to Churchill et al. [18], we decomposed each transformation matrix into translational and rotational components, which reveal the translation (mm) along and rotation (degrees) around the three axes.

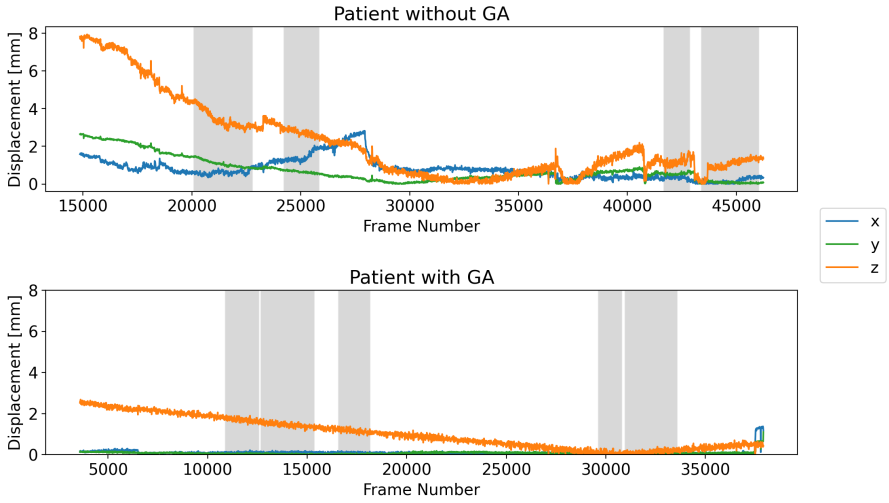


Figure 3: Example of motion curves for children with and without GA. Grey background inidcates the times where MR sequences have been acquired.

2.5. Statistical analysis

Statistical differences in motion metrics as well as in absolute translational and rotational components between groups and across brain regions (N=16) were evaluated using the Mann-Whitney U test. Correction for multiple comparisons across regions was carried out using False-Discovery Rate [23] (FDR), at FDR= 0.05.

3. Results

3.1. Exemplary motion curves

Exemplary motion curves during the entire scan session are shown in Figure 3 for one patient with and one without GA. Curves for other GA patients are similar to the shown example, whereas the curves for awake children show more variation.

3.2. Comparison of motion metrics

Figure 4 compares mean, median and maximum displacement, as well as motion free time for patients with and without GA. All metrics show a statis-

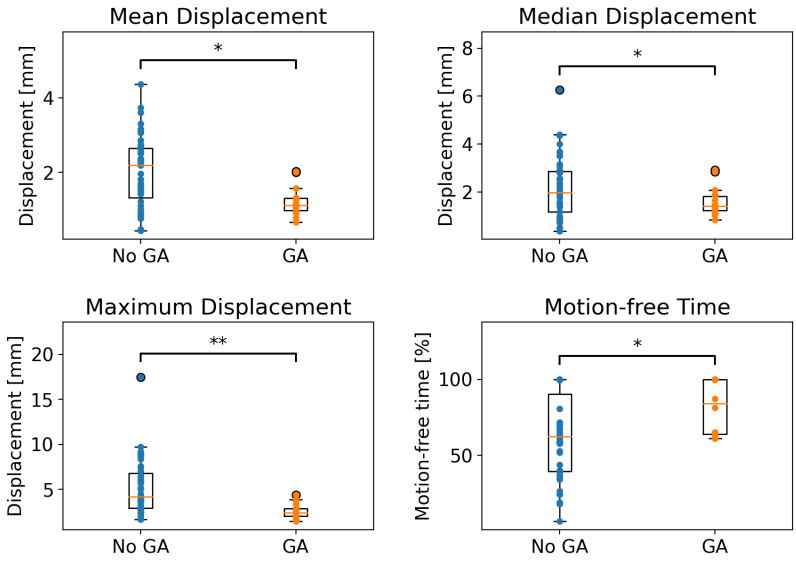


Figure 4: Comparison of motion metrics calculated with point cloud centroid for children with and without GA. Statistical significance after FDR correction is indicated by */** ($p < 0.05/p < 0.001$).

tically significant difference between anaesthetised and awake children. Please note several values corresponding to a substantial amount of motion even for children under anaesthesia.

The metrics for the 16 analysed brain regions can be compared in figures S.4, S.5, S.6 and S.7 of the Supplementary Material. Maximum and mean displacement as well as motion free time vary statistically significantly between patients with and without GA for all regions. Median displacements are distributed over a larger range and only show statistical significance for some regions, but the same trend as for the other metrics is still observable.

3.3. Decomposition of motion into translation and rotation

We extracted translational and rotational components for both patient groups and compare them in Figure 5. For both groups, the absolute translational component along the z-axis, which is the axis going into the scanner bore and pointing towards the top of the head, is significantly larger than the absolute

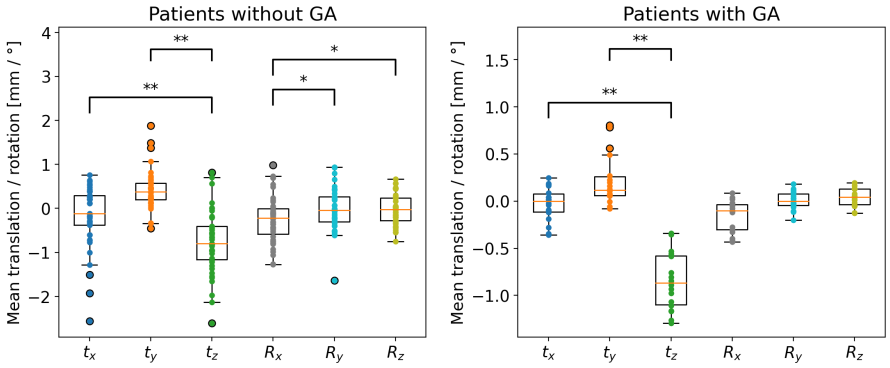


Figure 5: Decomposition of transformation matrices into translational (t_x, t_y, t_z) and rotational components (R_x, R_y, R_z), for children without and with GA. Statistical significance after FDR correction is indicated by */** ($p < 0.05/p < 0.001$); please note that significance was tested on absolute translations and rotations, whereas mean values of positive and negative components are shown in the graph for visualising the direction of the motion. Notice the different scales for children without and with GA.

values along the x- and y-axis. Regarding directionality of this translation, both patient groups move in the negative z-direction, i.e. slide downwards out of the scanner. For the children without GA, the absolute rotational component around the x-axis is significantly larger than the components around the other axes, whereas for the GA group no significant differences between the rotations are observed. In the RAS system, the x-axis points from right to left.

3.4. Motion of different brain regions

To further analyse the motion patterns, we calculated the median displacement on the x-, y- and z-axis for each brain region, respectively. Figure 6 shows the results for 4 cortical and 4 subcortical regions. The results for the remaining 8 regions analysed in this study are available in the Supplementary Material (Figure S.8). The magnitude of the motion on the z-axis exceed that of the other axes for both patient groups, even though the differences are smaller for the group without GA. This difference is statistically significant for both groups and all 16 regions except the comparison of z- and y-axis for the right hemisphere precentral region in the patient group without GA (see Figure S.8).

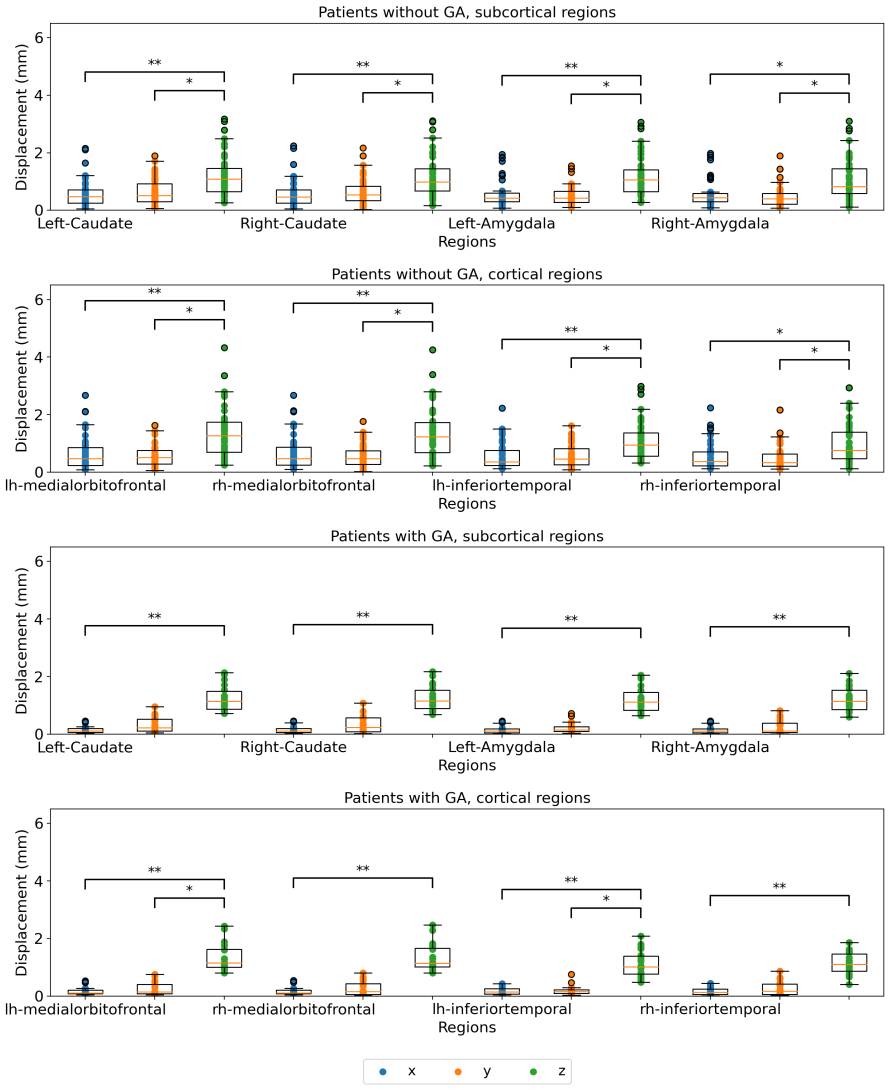


Figure 6: Comparison of median distance on each axis for 4 cortical and 4 subcortical regions split for the x-, y, and z-axis (blue, orange and green, respectively). The upper two plots show patients without GA, the lower two patients with GA. Statistical significance after FDR correction is indicated by */** ($p < 0.05/p < 0.001$).

4. Discussion

We analysed motion data from anaesthetised and awake children during MRI examination using the same protocol and the same MR scanner for all scans, which to our knowledge has not been reported so far.

Comparing motion metrics like mean, median, maximum displacement and motion-free time in Figure 4, confirmed our expectations that children without GA moved significantly more than children with GA for all analysed metrics. Nevertheless, residual movement was still observed for anaesthetised children. As described in the Introduction, Afacan et al. [1] investigated the correlation of motion metrics like mean and maximum displacement and motion-free time with image quality based on motion estimates, which were measured with two sensors placed on the patient's forehead. For enabling comparisons with their work, we calculated our metrics for the point cloud centroid, which corresponds to a point close to the middle of the nose bridge, as visualised in Supplementary Figure S.3. We reported both mean and median displacement, since the displacements over time were not normally distributed, which we confirmed with a Kolmogorow-Smirnow test for normality. Even though Afacan et al. did not find significant differences between radiologic evaluation and maximum displacement, we reported maximum movement since it provides valuable information about the severity of motion spikes, especially in the context of motion correction.

Furthermore, in Figure 4 we showed that even anaesthetised children moved substantially across all metrics: with mean displacements of up to 1.9 mm, maximum displacements of up to 4.4 mm and motion-free times down to 58 %. Afacan et al. [1], who analysed the correlation of motion metrics with radiologic evaluation of image quality, reported motion-free time medians of 38 %±21, 74 %±27, 97 %±17 and 99 %±3 for the grades 1, 2, 3, and 4, respectively (4 being the best image quality). Together with our results of motion-free time values down to 58 %, this indicates that GA does not guarantee high image quality due to considerable residual motion in several cases. Our different

threshold for calculating motion-free times, does not falsify these conclusions, since lowering our threshold from 2 mm to 0.2 mm (as in [1]) would only decrease our percentages of motion free times. Please note that Afacan et al. also used a different motion tracking device, an electromagnetic tracker [24] developed by Robin Medical Inc. (Baltimore, MD). For other motion tracking devices using external fitted tools, like the Polaris Vicra, inaccuracy in the motion estimates has been shown due to dislocation of the tool relative to the head [25].

In order to obtain more insight into which types of movement children perform most commonly during MR examinations, we decomposed the transformation matrices for each time point into translation and rotation. For both patient groups, we compared the components along / around the three axes. In Figure 5, we showed for both patient groups significantly larger absolute translation along the z-axis compared to x- and y-axes. For the children without GA, we additionally observed significantly larger rotations around the x-axis compared to the y- and z-axes. The z-translation, which corresponds to translation along the axis through the scanner bore, can be interpreted as a drift motion, especially for the children with GA. This becomes apparent when looking at the complete motion data throughout the whole scan session e.g. in Figure 3. For most of the GA patients one can observe a continuous increase in the displacement along the z-axis from the beginning until the end of the examination. Since we observed negative z-translation values for both patient groups, this motion corresponds to a gliding downwards out of the scanner. The rotational component around the x-axis observed for children without GA corresponds to a nodding motion, which together with a sliding in negative z-direction has previously also been observed as primary motion for adults in the MR scanner [18, 26]. For children specifically, this movement can be explained by parents or a screen with a movie being positioned at the end of the scanner bore. Thus, our analysis provides important conclusions for clinical examinations of children, namely that avoiding nodding motion should play a larger role in training children before an examination.

Finally, we analysed how the children's motion affects different parts of the

brain. For this, we chose 8 cortical and 8 subcortical regions of the brain, including paired regions from the left and right side of the brain. This allowed us to compare the influence of motion on the left and right side as well as back, front and central parts of the brain. We calculated the displacement along each axis for the selected regions. All in all, this analysis is in accordance with the results from the matrix decomposition, i.e. that the largest displacement was observed for the z-axis (cf. Figure 6). For the GA group very little motion was observed on the x- and y-axes. The group without GA showed more motion on the x- and y-axes, but the z-axis was still dominating. The differences were statistically significant for all regions, except the precentral region of the right hemisphere for the patient group wihout GA. For this region, the median motion was more distributed across y- and z-axes.

Our study is not without limitations. First, it is limited by the unbalanced distribution between the groups with and without GA. A larger GA group might enable higher statistical significance for comparing the motion metrics between both groups. Another potential source of error is the fact that FreeSurfer is optimised for adult brain anatomy, which could lead to slightly wrong segmentation of the brain regions. However, each segmentation was verified manually. Lastly, we excluded approximately one third of the data set, amongst other things due to poor scan quality impeding a successful segmentation by FreeSurfer. This could lead to excluding scans with large amounts of motion and thus, underestimating the true motion of children in general. However, our comparison of included and excluded scans in Supplementary Figure S.2 confirmed that excluded scans did not systematically have larger amounts of motion, apart from two outlier scans (see Figure S.2).

5. Conclusion

Considering the presented results together with the discussed limitations, our study puts some light onto how children move during a (PET/)MR scan, when they are awake as well as anaesthetised. First of all, we showed that even

children under GA show substantial amount of motion, which strengthens the need for alternative methods such as motion correction techniques and adequate preparation of the children to avoid motion artefacts in case the motion causes clinically image degrading artefacts. For the anaesthetised children, the clinical evaluation of image quality reported in this study shows no problem with motion artefacts; free motion, however, requires motion correction to avoid GA for more children. In addition, our data indicates that clinical routines as well as training methods under development should be given special attention to prevent nodding motion. Similarly, the higher prevalence of nodding motion and translation along the z-axis should be taken into account when optimizing and testing motion correction methods. In order to limit motion artefacts further, the application of MR sequences that are more robust towards these two types of motion regarding slicing and phase encoding direction should be considered.

Acknowledgment

We highly appreciate the scanner assistance of technologists Karin Stahr and Marianne Federspiel, and of radiographers Jákup Martin Poulsen and Nadia Azizi. This work was supported by The Danish Childhood Cancer Foundation (2014-34, 2015-48), as well as the Elsass Fonden (18-3-0147) and an unrestricted grant from the Novo Nordisk and Novozymes Talent Program.

Data and Code availability statement

Due to regulations according to the Danish Data Protection Agency, the very sensitive patient data including the MRI sequences are not available.

The code used to preprocess and create the results used in this manuscript mostly consists of open-source software (FreeSurfer and Python). Code for running the analysis can be found at GitHub.

References

[1] O. Afacan, B. Erem, D. P. Roby, N. Roth, A. Roth, S. P. Prabhu, S. K. Warfield, Evaluation of motion and its effect on brain magnetic resonance image quality in children, Pediatric Radiology 46 (12) (2016) 1728–1735. doi:10.1007/s00247-016-3677-9.

[2] F. Godenschweger, U. Kägebein, D. Stucht, U. Yarach, A. Sciarra, R. Yakupov, F. Lüsebrink, P. Schulze, O. Speck, Motion correction in MRI of the brain, Physics in Medicine and Biology 61 (5) (2016) R32–R56. doi:10.1088/0031-9155/61/5/R32.

[3] J. B. Andre, B. W. Bresnahan, M. Mossa-Basha, M. N. Hoff, C. Patrick Smith, Y. Anzai, W. A. Cohen, Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations, Journal of the American College of Radiology 12 (7) (2015) 689–695. doi:10.1016/j.jacr.2015.03.007.
URL http://dx.doi.org/10.1016/j.jacr.2015.03.007

[4] H. M. De Bie, M. Boersma, M. P. Wattjes, S. Adriaanse, R. J. Vermeulen, K. J. Oostrom, J. Huisman, D. J. Veltman, H. A. Delemarre-Van De Waal, Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans, European Journal of Pediatrics 169 (9) (2010) 1079–1085. doi:10.1007/s00431-010-1181-z.

[5] S. B. Runge, N. L. Christensen, K. Jensen, I. E. Jensen, Children centered care: Minimizing the need for anesthesia with a multi-faceted concept for MRI in children aged 4–6, European Journal of Radiology 107 (July) (2018) 183–187. doi:10.1016/j.ejrad.2018.08.026.
URL https://doi.org/10.1016/j.ejrad.2018.08.026

[6] J. M. Slipsager, S. L. Glimberg, J. Søgaard, R. R. Paulsen, H. H. Johannesen, P. C. Martens, A. Seth, L. Marner, O. M. Henriksen, O. V. Olesen, L. Højgaard, Quantifying the Financial Savings of Motion Correction in

Brain MRI: A Model-Based Estimate of the Costs Arising From Patient Head Motion and Potential Savings From Implementation of Motion Correction, Journal of Magnetic Resonance Imaging 52 (3) (2020) 731–738. doi:10.1002/jmri.27112.

[7] J. E. Havidich, M. Beach, S. F. Dierdorf, T. Onega, G. Suresh, J. P. Cravero, Preterm versus term children: Analysis of sedation/anesthesia adverse events and longitudinal risk, Pediatrics 137 (3). doi:10.1542/peds.2015-0463.

[8] K. Johnson, A. Page, H. Williams, E. Wassemer, W. Whitehouse, The use of melatonin as an alternative to sedation in uncooperative children undergoing an MRI examination, Clinical Radiology 57 (6) (2002) 502–506. doi:10.1053/crad.2001.0923.

[9] E. Törnqvist, Å. Månsson, I. Hallström, Children having magnetic resonance imaging: A preparatory storybook and audio/visual media are preferable to anesthesia or deep sedation, Journal of Child Health Care 19 (3) (2015) 359–369. doi:10.1177/1367493513518374.

[10] A. D. Edwards, O. J. Arthurs, Paediatric MRI under sedation: Is it necessary? What is the evidence for the alternatives?, Pediatric Radiology 41 (11) (2011) 1353–1364. doi:10.1007/s00247-011-2147-7.

[11] R. Frost, P. Wighton, F. I. Karahanoğlu, R. L. Robertson, P. E. Grant, B. Fischl, M. D. Tisdall, A. van der Kouwe, Markerless high-frequency prospective motion correction for neuroanatomical MRI, Magnetic Resonance in Medicine 82 (1) (2019) 126–144. doi:10.1002/mrm.27705.

[12] M. D. Tisdall, A. T. Hess, M. Reuter, E. M. Meintjes, B. Fischl, A. J. Van Der Kouwe, Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI, Magnetic Resonance in Medicine 68 (2) (2012) 389–399. doi:10.1002/mrm.23228.

[13] J. Maclaren, M. Herbst, O. Speck, M. Zaitsev, Prospective motion correction in brain imaging: A review, Magnetic Resonance in Medicine 69 (3) (2013) 621–636. doi:10.1002/mrm.24314.

[14] J. M. Slipsager, S. L. Glimberg, D. Gallichan, L. Højgaard, R. R. Paulsen, O. V. Olesen, Clinical Implementation of Fast Markerless Motion Correction in K-Space of Structural 3D MR-Images of the Brain, in: ISMRM 2019, 2019.

[15] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, J. C. Ye, K-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI, Magnetic Resonance in Medicine 61 (1) (2009) 103–116. doi:10.1002/mrm.21757.

[16] J. Zbontar, F. Knoll, A. Sriram, T. Murrell, Z. Huang, M. J. Muckley, A. Defazio, R. Stern, P. Johnson, M. Bruno, M. Parente, K. J. Geras, J. Katsnelson, H. Chandarana, Z. Zhang, M. Drozdzal, A. Romero, M. Rabbat, P. Vincent, N. Yakubova, J. Pinkerton, D. Wang, E. Owens, C. L. Zitnick, M. P. Recht, D. K. Sodickson, Y. W. Lui, fastMRI: An Open Dataset and Benchmarks for Accelerated MRI (2018) 1–35arXiv:1811.08839.
URL http://arxiv.org/abs/1811.08839

[17] H. Zhang, B. Dong, A Review on Deep Learning in Medical Image ReconstructionarXiv:1906.10643.
URL http://arxiv.org/abs/1906.10643

[18] N. W. Churchill, A. Oder, H. Abdi, F. Tam, W. Lee, C. Thomas, J. E. Ween, S. J. Graham, S. C. Strother, Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. Standard temporal motion and physiological noise correction methods, Human Brain Mapping 33 (3) (2012) 609–627. doi:10.1002/hbm.21238.

[19] L. Marner, K. Nysom, A. Sehested, L. Borgwardt, R. Mathiasen, O. M. Henriksen, M. Lundemann, P. Munck af Rosenschöld, C. Thomsen,

L. Bøgeskov, J. Skjøth-Rasmussen, M. Juhler, A. Kruse, H. Broholm, D. Scheie, T. Lauritsen, J. L. Forman, P. S. Wehner, L. Højgaard, I. Law, Early postoperative 18f-fet pet/mri for pediatric brain and spinal cord tumors, Journal of Nuclear Medicine 60 (8) (2019) 1053–1058. arXiv:https://jnm.snmjournals.org/content/60/8/1053.full.pdf, doi:10.2967/jnumed.118.220293.
URL https://jnm.snmjournals.org/content/60/8/1053

[20] B. Fischl, FreeSurfer, NeuroImage 62 (2) (2012) 774–781. doi:10.1016/j.neuroimage.2012.01.021.

[21] J. M. Slipsager, A. H. Ellegaard, S. L. Glimberg, R. R. Paulsen, M. Dylan Tisdall, P. Wighton, A. Van Der Kouwe, L. Marner, O. M. Henriksen, I. Law, O. V. Olesen, Markerless motion tracking and correction for PET, MRI, and simultaneous PET/MRI, PLoS ONE 14 (4). doi:10.1371/journal.pone.0215524.

[22] O. V. Olesen, M. R. Jørgensen, R. R. Paulsen, L. Højgaard, B. Roed, R. Larsen, Structured Light 3D Tracking System for Measuring Motions in PET Brain Imaging, in: Proceedings of SPIE Medial Imaging - The International Society for Optical Engineering, Vol. 7625, 2010, pp. 121–135.

[23] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing, Journal of the Royal Statistical Society: Series B (Methodological) 57 (1) (1995) 289–300. arXiv:https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1995.tb02031.x, doi:https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1995.tb02031.x

[24] A. Gholipour, M. Polak, A. van der Kouwe, E. Nevo, S. K. Warfield, Motion-robust mri through real-time motion tracking and retrospective

super-resolution volume reconstruction, Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2011 (2011) 5722—5725. doi:10.1109/iembs.2011.6091385.
URL https://europepmc.org/articles/PMC3687083

[25] S. H. Keller, M. Sibomana, O. V. Olesen, C. Svarer, S. Holm, F. L. Andersen, L. Højgaard, Methods for motion correction evaluation using 18f-fdg human brain scans on a high-resolution pet scanner, Journal of Nuclear MedicinearXiv:https://jnm.snmjournals.org/content/early/2012/02/13/jnumed.111.095240.full.pdf, doi:10.2967/jnumed.111.095240.
URL https://jnm.snmjournals.org/content/early/2012/02/13/jnumed.111.095240

[26] S. H. Keller, C. Hansen, C. Hansen, F. L. Andersen, C. Ladefoged, C. Svarer, A. Kjær, L. Højgaard, I. Law, O. M. Henriksen, A. E. Hansen, Motion correction in simultaneous PET/MR brain imaging using sparsely sampled MR navigators: a clinically feasible tool, EJNMMI Physics 2 (1) (2015) 1–7. doi:10.1186/s40658-015-0118-z.
URL http://dx.doi.org/10.1186/s40658-015-0118-z

Supplementary Material

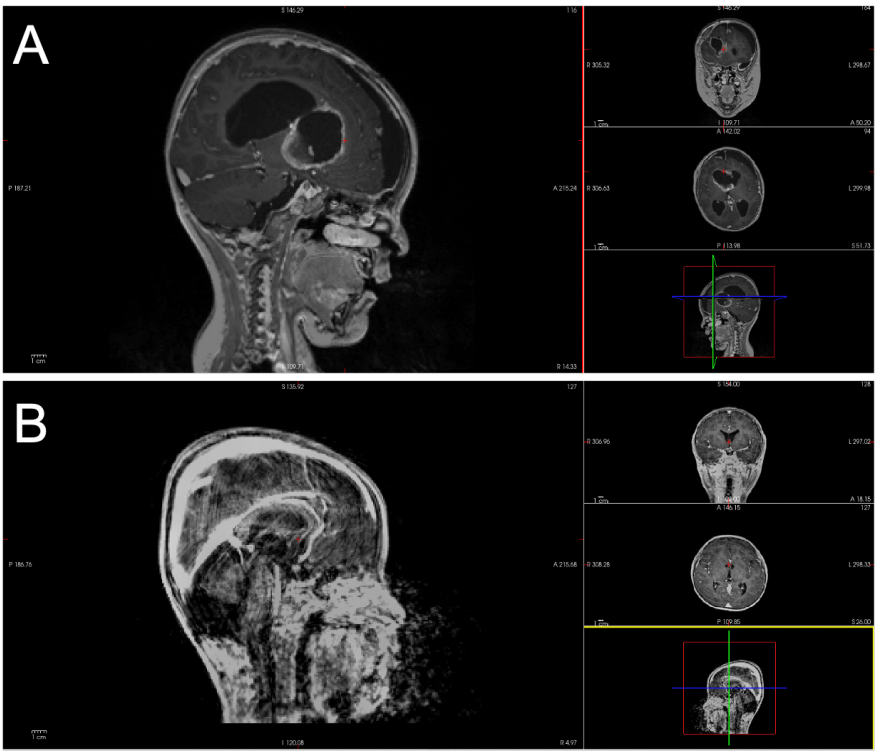


Figure S.1: Examples of excluded scans due to (A) large defects and (B) poor scan quality.

Table S.1: Parameters of the MR sequences describing the type of acquisition, flip angle, repetition, echo as well as inversion time and voxel size.

Sequence	**Acquisition type**	**Flip angle [°]**	**TR/TE/TI [ms]**	**Voxel size [mm]**
T_1-weighted MPRAGE	3D	9	1900/2.52/900	1.00×1.00×1.00
T_1-weighted STIR	2D	150	2000/34/800	0.45×0.45×4.00
T_2-weighted FLAIR (transversal)	2D	150	9000/85/2500	0.69×0.69×4.00
T_2-weighted FLAIR (coronal)	2D	130	9000/95/2500	0.43×0.43×4.00
T_2-weighted Blade	2D	140	4000/118/—	0.72×0.72×5.00

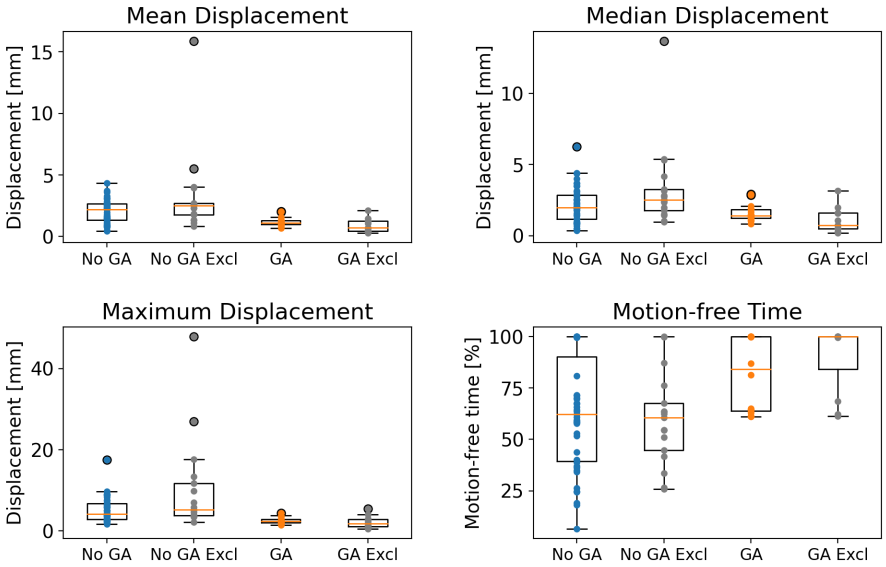


Figure S.2: Motion metrics for point centroid of scans excluded due to poor image quality compared to included scans. The analysis of the excluded scans is performed analogously to the analysis shown in Figure 4. Except two outlier scans with significantly larger motion in the group without GA, the excluded patients show similar amount of motion as the included patients.

Figure S.3: Example of the Tracoline point cloud for an adult. The point cloud centroid is schematically visualised by the orange circle. Please note that the field of view varies with head size and curvature of the face; for children a larger part of the face might be covered.

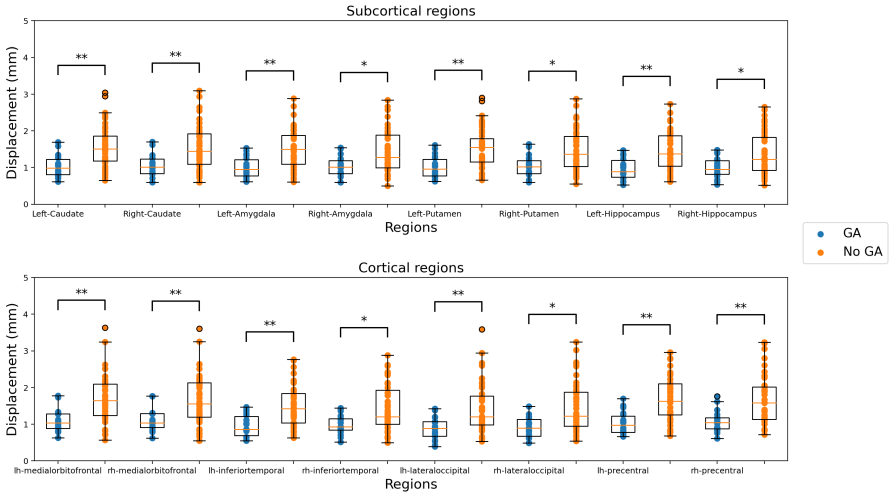


Figure S.4: Comparison of mean displacement for all cortical and subcortical regions. Statistical significant differences after Benjamini-Hochberg correction are indicated by */** ($p < 0.05/p < 0.001$).

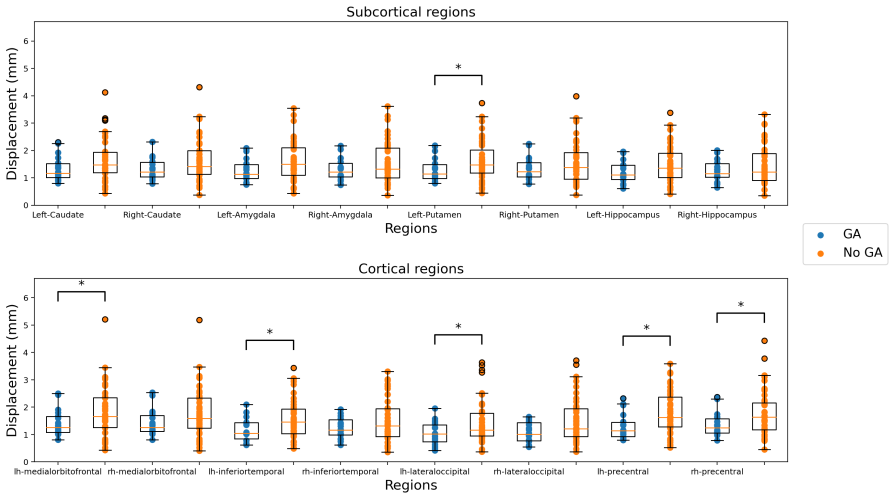


Figure S.5: Comparison of median displacement for all cortical and subcortical regions. Statistical significant differences after Benjamini-Hochberg correction are indicated by * ($p < 0.05$).

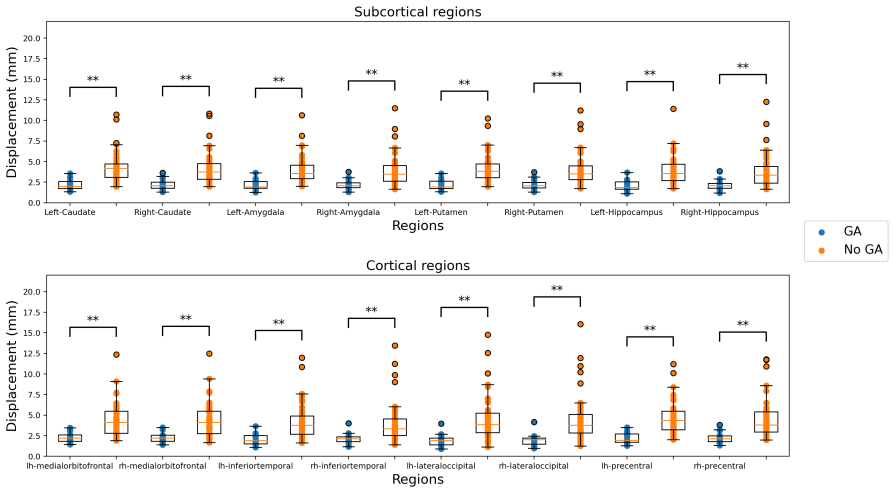


Figure S.6: Comparison of maximum displacement for all cortical and subcortical regions. Statistical significance after Benjamini-Hochberg correction is indicated by ** ($p < 0.001$).

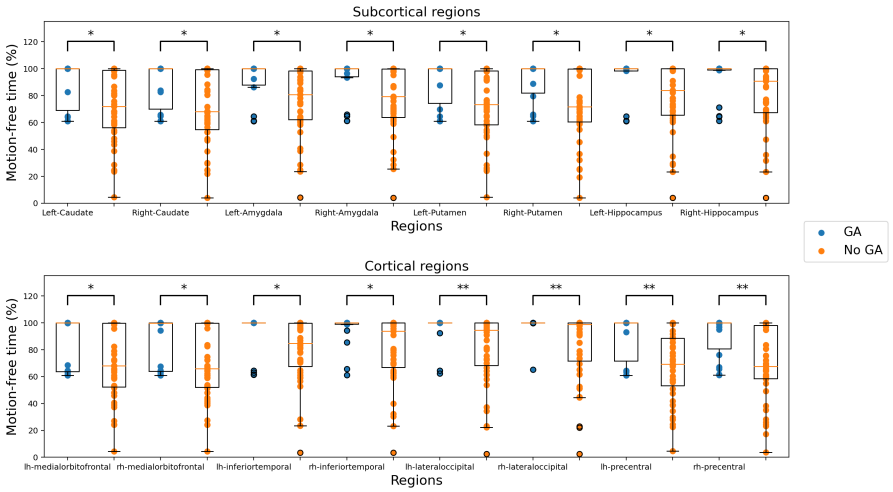


Figure S.7: Comparison of motion-free time for all cortical and subcortical regions. Statistical significance after Benjamini-Hochberg correction is indicated by */** ($p < 0.05/p < 0.001$).

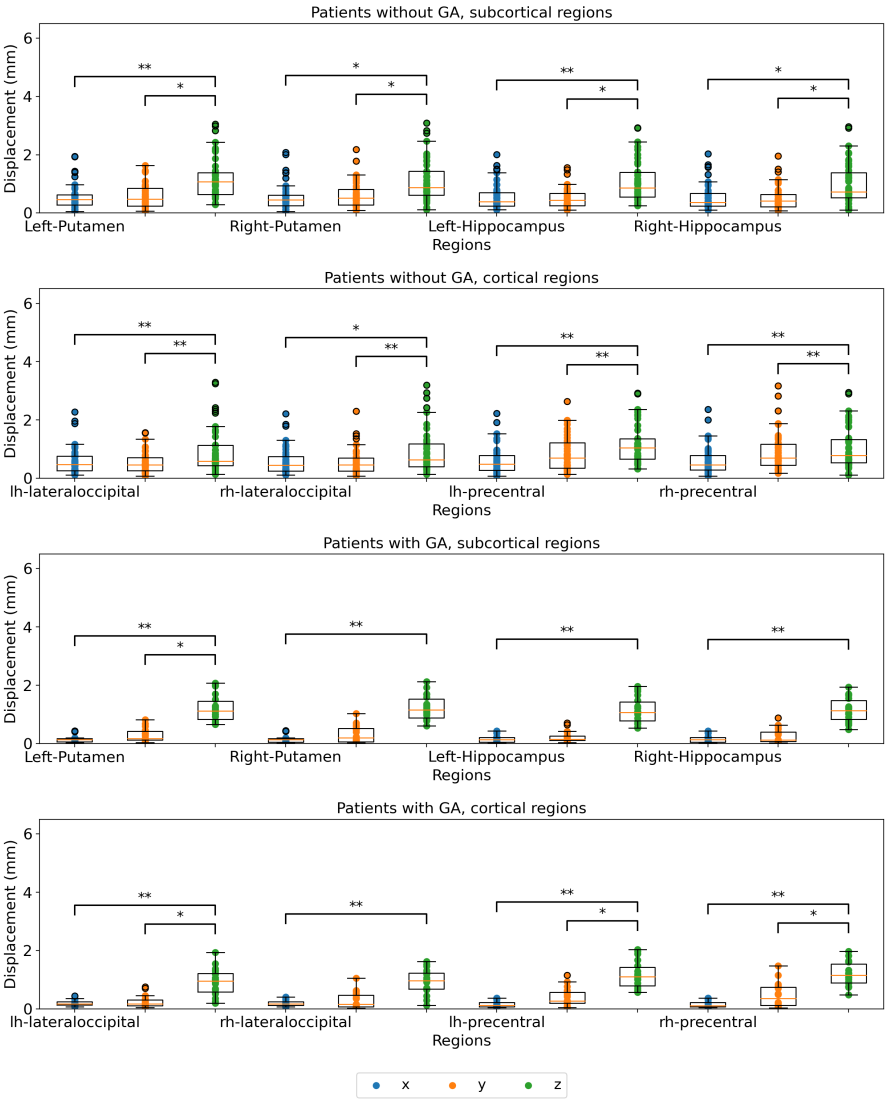


Figure S.8: Comparison of median distance on each axis for the 4 cortical and 4 subcortical regions not included in Figure 6. The top two plots show patients without GA, the lower two patients with GA. Statistical significance after Benjamini-Hochberg correction is indicated by */** ($p < 0.05/p < 0.001$).