Validation of Quantitative MRI: Fat Quantification and ADC Mapping

In the Head-and-Neck Area

Master Thesis

Renske de Jong

2021

TU Delft

LUMC

Cover image: Art work by Angela Palmer.
Engraved Self Portrait (on 14 sheets of Mirogard glass) based on MRI scans of the artist.

Validation of quantitative MRI: fat quantification and ADC mapping

In the head-and-neck area

by

Renske Maria de Jong

to obtain the degree of

Master of Science
in Biomedical Engineering

at the Delft University of Technology,
to be defended publicly on Tuesday November 9th, 2021 at 2:00 PM.

Student number:	4496671	
Project duration:	January 18th, 2021 – November 9th, 2021	
Thesis committee:	Dr. F. M. Vos,	TU Delft (Chair)
	Dr. E. Astreinidou,	LUMC (Daily supervisor)
	Dr. ir. M. Staring,	TU Delft/LUMC

An electronic version of this thesis will be available at `http://repository.tudelft.nl/`.

TU Delft

Preface

Almost 10 months ago, I started this project. It was something I have been both dreading and looking forward to for years. However, I did not imagine the project to run in these circumstances, where working from home would be 'the new normal'. Even though it has not been most optimal for me, thankfully I have received help from many dear people around me.

Firstly, I would like to say thank you to my daily supervisor, Eleftheria. COVID-19 has had an enormous impact on everyone, including the LUMC of course, but I have really appreciated your effort to give me as much opportunities during the last months as possible. Not only did you assist me professionally, but you were also very understanding on a more personal level. Ευχαριστώ πολύ!

Filipa, thank you for helping me put everything into perspective whenever I felt overwhelmed in the process. The stories on your experiences stimulated me to proceed in good spirits. Also, thank you for all your work in the scanning sessions and your super quick responses to my emails! It was very nice working with you. Muito obrigado!

Frans, thank you in the first place for conveying your enthusiasm about the field of medical physics. It has inspired me so much, that I had no doubts about which specialization of the BME Master program was my favorite. Although to be fair, I did have some doubts about my background and abilities along the way, but you have always been available to offer me help and advise whenever I reached out to you. Heel erg bedankt!

Thanks to everyone from the LUMC C.J. Gorter Center for High Field MRI and the Antoni van Leeuwenhoek Hospital who helped me out and supplied me with some necessary items for my research. Aashley, Wouter, Hermien, Petra, Uulke, Robin, Ernst, it was fun learning from you, I appreciate all your help!

Kevin and Nick, thank you in particular for all your time and effort that went into explanations of theory and assistance during experiments. It has been fun working with you both! You both have shared great advice and gave me some insight into what research is really like.

Thanks to all my friends for supporting me and hearing me out whenever I was feeling excited or stressed. Marjolijn, you are the sweetest. My hockey buddies from Hudito Dames 2, Blub, my roomies from Delft and my new Rotterdam roomies, thank you for being there!

Lastly, thank you Papa, Mama, and Sjoerd, for your boundless support throughout these last couple of months and all the other years of my academic career. Starting the Biomedical Engineering Master program after obtaining my Bachelor's degree in Industrial Design Engineering has been a challenge. Some things have - to my frustration - taken much effort and mostly time to understand. But I have learnt that knowledge and experience just don't come to you overnight; like Rome wasn't built in a day. This thesis marks the end of my academic education and the beginning of a whole new phase. And I am excited to find out what's next!

Abstract

Purpose. Fat fraction (FF) and apparent diffusion coefficient (ADC) values estimated by Dixon MRI and diffusion weighted MRI (DWI) techniques respectively, are relatively new quantitative imaging parameters and increasingly accepted as imaging biomakers for all sorts of purposes. The aim of the BOCASEcA study is to research whether these techniques can be used as biomarkers for patient-reported xerostomia and dysphagia post-radiotherapy. In this project, steps have been taken to validate the use of certain fat quantification and ADC mapping protocols in the BOCASEcA study.

Method. mDIXON Quant is a Philips product designed for MR fat quantification. We performed phantom studies and a healthy volunteer study to evaluate the accuracy and repeatability of a standard mDIXON Quant protocol with default parameters and an mDIXON Quant protocol that is used in LUMC on muscles throughout the whole body. Another phantom study was done to evaluate the (geometrical) accuracy of DWI-SPLICE, a technique that can be used for ADC mapping. This DWI technique is known to have less susceptibility issues than conventional EPI-DWI. We tested both the accuracy and deforming artifacts for an EPI sequence and a clinically used DWI-SPLICE protocol from LUMC.

Results. Adequate accuracy and robustness were observed for the standard Philips mDIXON Quant protocol. The LUMC muscle protocol, however, yielded incorrect measurements that were underestimations of the real FF values. The DWI-SPLICE protocol showed better geometrical accuracy than the EPI protocol. Accuracy of the ADC measurements was sufficient for ADC values higher than 0.6 x 10^{-3} mm^2/s which is a clinically relevant range.

Discussion. Since the accuracy of both the standard Philips mDIXON Quant protocol and DWI-SPLICE protocol was validated, it is recommended that they are used and further optimized for the BOCASEcA study.

Contents

1 Introduction

The LUMC department of radiotherapy is working closely together with the departments of otorhinolaryngology and radiology and with the Holland Proton Therapy Center (HPTC) and the Antoni van Leeuwenhoek (AVL) on a project called *BOCASEcA: MR imaging biomarkers associated with patient-reported xerostomia and dysphagia post-proton beam therapy*. The purpose of the BOCASEcA study is to identify MRI biomarkers using quantitative MRI techniques, to reveal damage to tissues involved in saliva production and in swallowing for head-and-neck cancer patients after radiotherapy treatment. The ultimate aim is to reduce toxicity and improve treatment planning for HNC radiotherapy. This thesis report presents the work that is done to validate the performance of two quantitative MRI techniques that have the potential to identify promising biomarkers. In this first chapter, the context of the thesis project is introduced.

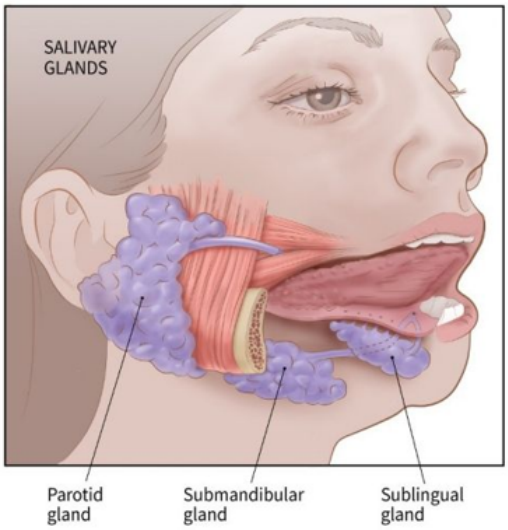
1.1. Head-and-Neck Cancer

Head-and-neck cancer (HNC) is a term that covers a broad range of squamous cell carcinomas in the upper aerodigestive epithelium - including the nasal cavity, oral cavity, paranasal sinuses, pharynx and larynx. Risk factors such as tobacco, alcohol and the human papillomavirus (HPV) can contribute to the development of these epithelial malignancies [1, 2]. The incident rate of HNC in the Netherlands is approximately 3000 patients a year [3]. One of the most important treatment modalities for head-and-neck cancer is external radiotherapy. Unfortunately, radiotherapy in the head-and-neck region often causes serious side-effects such as xerostomia (dry mouth syndrome) and dysphagia (difficulty in swallowing). These radiation-induced pathologies can substantially impair the patient's quality of life (QoL).

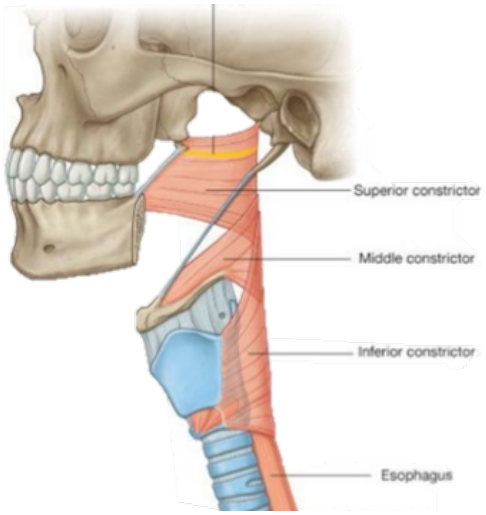
1.2. Radiation-Induced Xerostomia and Dysphagia

Some involved organs at risk (OARs) during radiotherapy are shown in Figure 1.1. The three organs that are responsible for about 90% of the total saliva production are visualized in the left image. About 65% of unstimulated saliva production is due to the submandibular gland, 20% comes from the parotid glands and about 7 - 8% from the sublingual gland. The contribution of the parotid glands increases to about 50% during stimulated saliva production [4]. The pharyngeal constrictor muscles (PCM), which are swallowing muscles, are shown on the right. Current IMRT recommendations focus on sparing the PCM and oral cavity.

Several publications have reported on the relationship between excessive radiation to the organs at risk (OARs) and two common HNC radiotherapy side-effects, xerostomia or dysphagia. The mean radiation dose to the parotid and submandibular glands is considered to be a confounding factor for xerostomia, while mean dose to the pharyngeal constrictor muscles (PCM) is considered as a confounding factor for dysphagia [5–8]. Improved treatment approaches, such as proton therapy, have the potential to reduce radiation dose to healthy tissues surrounding the tumor [9]. However, better insight in the dose-response relationship is desired to improve treatment planning.



(a) Salivary glands



(b) Pharyngeal constrictor muscles

Figure 1.1: Organs involved in saliva production (a) and primary muscles responsible for swallowing (b).

1.3. Quantitative MRI Parameter Mapping

Quantitative MRI (qMRI) techniques have expanded the potential to assess and quantify certain features of tissues that go beyond the possibilities of conventional MRI [10]. Furthermore, the superior soft tissue contrast resolution provided by MRI allows for great distinction of the similar soft tissues in the head-and-neck (e.g. glandular, muscular or tumorous tissues). Thus, using qMRI techniques in the head-and-neck, biomarkers can be found that reveal measurable radiation-induced damage to the salivary glands and swallowing muscles. Subsequently, the three dimensional image after radiotherapy can be correlated with the 3D dose distribution of the treatment plan in order to gain an actual spatial representation of the radiation-induced differences. Amongst promising qMRI techniques are multi-echo Dixon imaging, which relies on the difference in precession frequencies of water and fat, and diffusion weighted imaging (DWI), yielding the apparent diffusion coefficient (ADC).

1.3.1. MRI Fat Quantification

Dixon MR imaging is used as a tool to separate water and fat, which is useful e.g. for delineation of organs for radiotherapy planning. However, when signal is acquired using multiple echos and certain post-processing is applied, Dixon MRI has the potential to robustly quantify the fat content. The proportion of signal derived from fat protons, the fat fraction (FF), has emerged as an objective, image-based biomarker for illnesses in the last couple of years [11]. Scientific interest in fat quantification using magnetic resonance imaging (MRI) and spectroscopy (MRS) has grown over the last two decades (Figure 1.2).

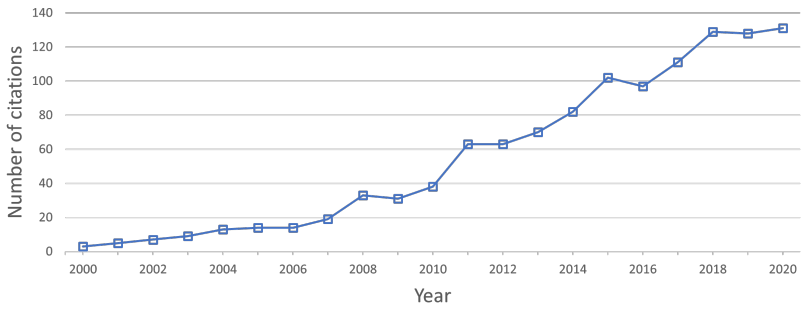


Figure 1.2: Number of publications found on PubMed per year for search term 'fat quantification magnetic resonance'.

During and after radiotherapy, the fat-to-water ratio in the OARs is expected to change, since the radiation dose will result in a loss of acinar cells [12, 13]. This suggests that the ratio of acinar and adipose cells is a measure for the radiation damage that induces xerostomia and dysphagia. This ratio can be macroscopically measured using Dixon imaging. However, there are limited publications that report Dixon MR measurements of fat content in the salivary glands and PCM. Table 1.1 shows various fat fraction values for both parotid and submandibular glands that are found in relevant studies that used quantitative Dixon imaging. Some of these studies have been done to either compare presumed healthy tissue to pathological tissue or to compare pre-radiation therapy treatment tissue to post-treatment tissue. Only the values acquired in a healthy control group or in HNC patients before treatment are given in the table.

Author, Year	N (#)	Age (years)	BMI	FF PG (%)	FF SMG (%)
Chang et al., 2013 [14]	114	46.5 ± 13.9	NR	46.0 ± 12.8	NR
Kise et al., 2017 [15]	16	37.6 ± 10.0	22.0 ± 3.3	30.7 ± 13.7	5.9 ± 4.9
Chikui et al., 2018 [16]	46	51.2 ± 19.6	22.3 ± 3.8	36.1 ± 11.4	8.9 ± 5.4
Zhou et al., 2018 [17]	41	49.1 ± 11.5	NR	38.2 ± 9.7	NR
Su et al., 2019 [18]	87	52.5 ± 14.8	NR	37.5 ± 10.5	12.9 ± 4.7
Chu et al., 2020 [19]	28	51.4 ± 9.8	21.8 ± 3.0	30.9 ± 8.3	4.7 ± 3.1

Table 1.1: Overview of fat fraction values for salivary gland tissue of healthy volunteers measured in previous studies. Some studies did not report BMI or research the SMG FF. N = number of subjects, BMI = body mass index, FF = fat fraction, PG = parotid gland, SMG = submandibular gland, NR = not reported.

Humbert et al. have measured the fat fraction in healthy tongue tissue using the quantitative Dixon technique from GE (GE IDEAL-FSE) and they found a fat fraction of 26.5 ± 3.5 (n = 10) [20]. For subcutaneous fat tissue, we expect a FF of 90% to 100% and for muscle tissue, we expect a FF of around 5%. However, muscle FF is strongly dependent on age.

1.3.2. Diffusion Weighted MR Imaging

Diffusion weighted imaging visualizes the Brownian motion of water molecules in tissues [21]. Two images are obtained in which the the first gradient dephases the water molecules and the second gradient rephases them. Stationary water molecules will yield approximately the same signal intensity before and after the gradients. However, moving water molecules will result in signal loss, because they will not be completely rephased by the second gradient. The amount of lost signal is proportional to the amount of movement [22]. The apparent diffusion coefficient (ADC) is a measure of the magnitude of the diffusion of water molecules within tissue [23].

In 2007, the Radiological Society of North America (RSNA) organized a collaboration initiative to unite researchers, healthcare professionals and the industry with the collective goal of developing radiology into a more quantitative science. The resulting group is called the Quantitative Imaging Biomarkers Alliance (QIBA). QIBA endeavors to develop hard- and software standards that will lead to more consistent, accurate and reproducible quantitative results across all imaging sites and times [24]. They have designed several protocols and phantoms for optimizing acquisition parameters and quality assessment of various qMRI techniques. The DWI-MRI committee has developed such a standardized imaging procedure including a phantom which meets specific performance claims to achieve reproducible quantitative measurements.

1.4. Purpose of this thesis

The purpose of this thesis is to validate the performance of two qMRI techniques; a fat quantification technique, the mDIXON Quant, and a diffusion weighted imaging technique, the DWI-SPLICE. Insight in their performance can be achieved by reviewing the accuracy and test-retest repeatability of the techniques. This study has two main objectives:

1. The first one is to calculate and verify the FF and ADC values of various MRI protocols using phantoms. This allows for assessment of the accuracy of the protocols. There are already commercial phantoms available to be used for the purpose of the study.

2. The second objective is to evaluate the repeatability of the qMRI protocols on phantoms and healthy volunteer data. This is necessary to validate the robustness of the protocols.

The outline of this thesis work is as follows. In Chapter 2, the two qMRI techniques are introduced and their working and relevance is explained. The approach that is used to achieve the main objectives and the corresponding experimental setup is given in Chapter 3. Subsequently, Chapter 4 includes the outcomes of the experiments. Lastly, in Chapter 5, the results are put into perspective, limitations are mentioned and recommendations are given.

2

Quantitative MRI Techniques

This Chapter elaborates on the background and working of mDIXON Quant and DWI-SPLICE.

2.1. mDIXON Quant

2.1.1. The basic Dixon method

The basic Dixon method was developed in 1984 by W. Thomas Dixon [25]. He introduced a method that allowed for the creation of separate fat and water images in MRI and thus the opportunity to suppress fat. The method is based on the phenomenon that water protons precess between 3 and 4 parts per million faster than fat protons [25]. In his work, Dixon created a multi-repetition single spin echo sequence for a 0.35 T scanner, in which he sampled two echoes with a slightly different echo delay. The frequency difference between fat and water molecules determines the relative timing for TE1 and TE2.

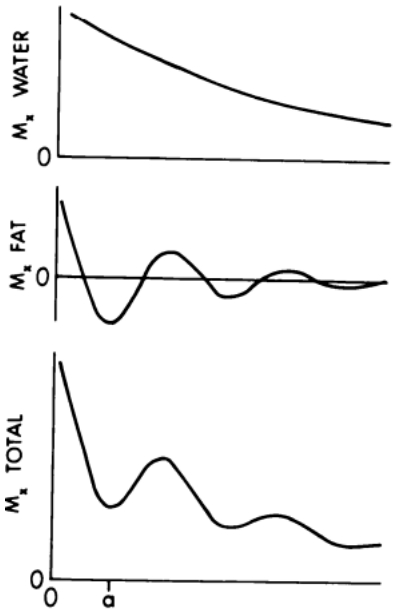
After a 90° pulse, all magnetization vectors point in the same direction. The water protons will then start to precess relatively faster than the fat protons. Since the water and fat protons were initially in phase, this is when the net magnetization was at its maximum. The first echo aligns the spins again after dephasing and generates an in-phase image. The second echo is then timed whenever the water and fat protons are exactly out of phase, at the magnetization minimum (see Figure 2.1a). A water-only and a fat-only image can be generated by summation and subtraction of the acquired images, respectively [26]. An example can be seen in Figure 2.1b. Not only does this allow for fat suppression - by using the water-only image-, but also for quantification of fat, calculated from the water-fat ratio. The resulted fat fraction (FF) is increasingly used as a tool to study and diagnose certain pathologies.

2.1.2. Quantitative Dixon imaging

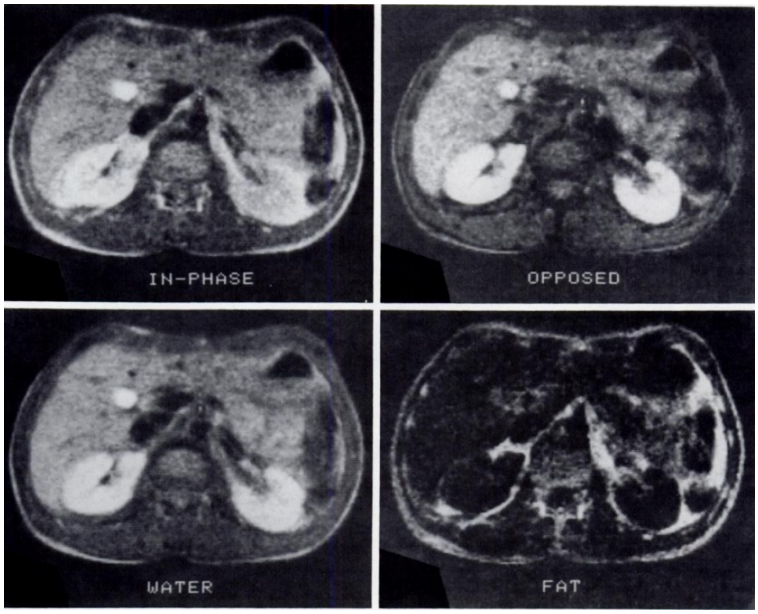
The basic Dixon method is sensitive to inhomogeneities in the magnetic field, which is an expanding problem while MRI field strengths are increasing. However, more robust fat fraction mapping is accomplished with the use of more echos and improved post-processing, i.e. iterative decomposition of water and fat with echo asymmetry and least-square estimation (IDEAL) [27]. This more reliable technique is frequently used in e.g. fatty infiltration in muscle tissue [10] and its outcome is often called fat fraction (FF) or proton density fat fraction (PDFF). Even though homogeneous fat saturation in the head-and-neck region is complicated to achieve, due to many air-tissue transitions [14], multi-echo Dixon imaging is able to provide stable measurements. Multiple MRI system vendors are implementing the IDEAL technique in their systems to determine fat contents, often optimized for liver tissue.

2.1.3. Philips mDIXON Quant

Philips designed *mDIXON Quant*, which makes use of multiple echos (at least six) during acquisition and a multipeak model, T2* correction and eddy current compensation during post-processing [28]. The higher



(a) Typical free induction decay; decay of magnetization in the X-direction in water, in fat and combined. Time *a* is when the first minimum occurs.



(b) Transverse image of a normal volunteer after a 20-hour fast. The in-phase and opposed phase images are shown on the top row. The water and fat images are shown below.

Figure 2.1: From the paper Simple Proton Spectroscopic Imaging by Thomas Dixon, 1984 [25]

number of echoes enables T2* correction, thus decreasing its influence on the relative signal intensity of the fat peaks. Separation of fat and water is done based on a spectral peak model, in which the precession frequencies of water and multiple types of fat are given. A 7-peak reconstruction enables better modeling of the heterogeneous fat distribution than 1- or 2-peak. A water and a fat image are constructed based on the complex raw data retrieved by the six echos. Subsequently, a parametric map with an FF value for each voxel (FF map) is automatically generated based on the resulting fat and water signal and using this equation:

$$I_{FF} = \begin{cases} \dfrac{|I_F|}{|I_F + I_W|}, & \text{if } I_F \geq I_W \\ 1 - \dfrac{|I_W|}{|I_W + I_F|}, & \text{if } I_F < I_W \end{cases} \tag{2.1}$$

Where I_F stands for fat intensity derived from the fat only image and I_W is water intensity derived from the water only image. Hu et al. (2021) proved that a six-echo Dixon protocol from all three mentioned vendors yielded accurate in-vitro proton density fat fraction values across imaging centers and field strengths [29].

The mDIXON Quant method is designed for fat quantification in the liver, however, the method can be applied to other regions as well. This is useful, since quantification of fat has important applications in various tissues throughout the entire body. During and after radiotherapy, the FF value in the salivary glands is expected to increase, since the radiation dose will result in a loss of acinar cells [12, 13]. This suggests that the ratio of acinar and adipose cells is a measure for the radiation damage that induces xerostomia and dysphagia. This ratio - i.e. the fat fraction (FF) - can be macroscopically measured using Dixon imaging. In a fat quantification study in myxoid licosarcomas (often located in extremities of the body) the ratio of fat to water was found to be increased after radiotherapy [30]. However, there is yet only one publication that reports dynamic changes of the fat content caused by radiotherapy in the head-and-neck region [17]. In this study, Zhou et al. found that the FF value in the parotid glands increased during radiotherapy and would then decrease to the original FF value after treatment was done.

2.2. DWI-SPLICE

2.2.1. Conventional EPI-DWI

DWI provides information on the local mobility of water in tissue. Dense structures, such as tumors but also glandular tissues, will show contrast with surrounding tissue. However, conventional DWI is not widely used in areas with many air-tissue interfaces such as the head-and-neck area, since strong susceptibility variations result in poor geometrical accuracy. The single-shot echo planar imaging (EPI) technique that is used in conventional DWI allows for very fast imaging and provides a good signal-to-noise ratio (SNR), but the geometrical image distortions can be multiple centimeters [31].

2.2.2. Philips DWI-SPLICE

A different technique that can be used for DWI is the single-shot turbo spin echo (TSE) technique. This technique is able to keep geometric information accurate, but at the cost of signal. One approach that successfully deals with the issues that come with TSE is SPLICE (split acquisition of fast spin echo signals for diffusion imaging). Using SPLICE, acquisition is done with both a spin echo and a stimulated echo. These echos are recorded separately - so-called split echo acquisition - and the images from the two echos are then reconstructed and merged, enabling the technique to produce higher SNR and better resolution than conventional TSE-DWI [31, 32].

3

Approach & Methodology

This thesis work includes the performance assessment of two qMRI techniques; mDIXON Quant and DWI-SPLICE. More elaborate work was done for the mDIXON Quant technique. This chapter describes the methodology and the experimental setup for the research that has been done. After an overview of the approach and an introduction of the protocols that are used, the experiments are elaborately described.

3.1. Approach

This project included experiments using MRI phantoms. Medical imaging phantoms are objects that are substitutes for human tissues. Phantom experiments allow for low-key measurements without ethical concerns and are a useful tool to assess repeatability and accuracy in a near ideal situation. The clinical usefulness and in-vivo robustness of the techniques were subsequently tested in MRI scanning sessions with healthy volunteers. An overview of the approach for performance evaluation of the two techniques can be seen in Figure 3.1. Due to the limited amount of time and the length of the DWI sequences, repeatability was not tested for the DWI-SPLICE.

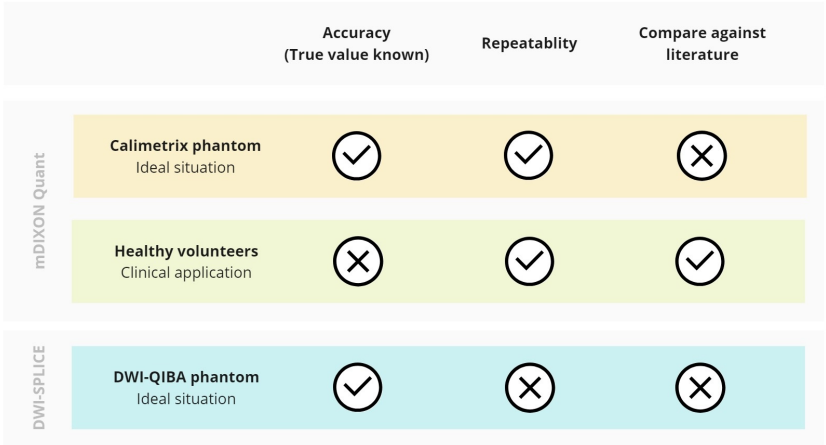


Figure 3.1: Overview of approach for performance evaluation of mDIXON Quant and DWI.

3.1.1. mDIXON Quant

Two mDIXON Quant protocols were applied on a MRI compatible fat quantification phantom to assess their performance based on accuracy. Subsequently, the best performing protocol was further investigated in a volunteer study. In Table 3.1, the two mDIXON Quant protocols are introduced. The first protocol was retrieved

from the LUMC C.J. Gorter Center, which from now on will be referred to as the 'LUMC muscle' protocol. This protocol is used for muscle MR imaging in the whole body. The other protocol is offered by Philips as the standard mDIXON Quant protocol with default system parameters (depending on the MRI system that is used), which will be referred to as the 'standard Philips' protocol.

In both protocols, a very low flip angle of 3° is used, in order to avoid T1 saturation. The standard protocol from Philips is optimized for liver imaging in one breath hold, which is why the sequence is very short (23 seconds against 6 minutes and 11 seconds for the LUMC muscle protocol). This is also the reason why the resolution is set relatively low. The actual repetition time, first echo time and echo spacing were 5.6 ms, 0.97 ms and 0.7 ms, respectively, for the scanners that we used (model: Philips Ingenia 3T). The LUMC muscle protocol is designed for a scanner with relatively low slew rate which is why the echo times are a bit longer.

	LUMC muscle	**Standard Philips**
Field of view (mm)	230 x 230 x 189	400 x 350 x 180
Acq. voxel size (mm)	1.4 x 1.4 x 3	2.5 x 2.5 x 3
Recon voxel size (mm)	1.03 x 1.03 x 3	2.08 x 2.08 x 1.5
Number of echos	6	6
First TE (ms)	2.27	'shortest'
ΔTE (ms)	1.52	'shortest'
TR (ms)	12	'shortest'
Flip angle (degrees)	3	3
Total scan duration	06:11.0	00:23.6
Act. TR/TE1/ΔTE (ms)	12 / 2.3 / 1.5	5.6 / 0.97 / 0.7
PNS / level	50% / normal	99% / 1st level
NSA	3	1

Table 3.1: Most relevant parameters of two protocols used in the experiments; the LUMC muscle and standard Philips protocol. Field of view and voxel size are given in (RL x AP x FH). TE = echo time, TR = repetition time, PNS = peripheral nerve stimulation, NSA = Number of Signals Averaged.

Based on the acquired data, an FF map is automatically calculated by a reconstruction method that Philips created. This method makes use of a multipeak fat spectral model and corrects for T2* decay and eddy current effects. The LUMC muscle protocol, however, is normally used in combination with an in-house reconstruction technique that is created by Kevin Keene and Aasley Sardjoe Mishre (LUMC C.J. Gorter Center for High Field MRI). This method also uses a multipeak fat spectral model and T2* correction, but no eddy current correction.

3.1.2. DWI-SPLICE

The current clinical pre-therapy MRI protocol for head-and-neck cancer patients in LUMC includes a DWI-SPLICE protocol. Because this sequence suffers less from susceptibility distortions and spatial blurring than conventional DW-imaging, it is appropriate for delineation of targets and OARs in radiotherapy treatment planning. To test the performance of the clinically used DWI-SPLICE protocol, we used a DWI phantom with known true ADC values. This thesis work does not include any DWI measurements on (healthy) human tissues.

The DWI-MR Biomarker committee of the QIBA created a DWI phantom and protocol combination to assess the performance of MR scanners. This QIBA protocol uses a single-shot EPI-DWI sequence with four b-values up to 2000. Both the clinical DWI-SPLICE used in LUMC and the EPI-DWI sequence from QIBA will be tested on its accuracy in this research and the outcomes will be compared. The most relevant parameters for the two DWI protocols (LUMC-SPLICE and QIBA-EPI) are given in Table 3.2.

The clinical protocol makes use of 2 b-values: 0 and 800 s/mm^2. However, the number of b-values chosen and their magnitude influence the resulting ADC values significantly [33]. Lower b-values reflect tissue perfusion, which results in higher ADC values, while higher b-values result in lower ADC values, since they reflect true diffusion. Using more lower b-values leads to a better estimation of perfusion effects, which in turn will allow for better estimation of the true diffusion. For this reason, an MRI protocol using DWI-SPLICE with four b-values of 0, 100, 200 and 800 s/mm^2 will be tested in this research.

	LUMC-SPLICE	QIBA-EPI
Field of view (mm)	280 x 228 x 108	220 x 220 x 124
ACQ voxel size (mm)	1.8 x 1.8 x 3	1.7 x 1.7 x 4
Recon voxel size (mm)	1.5 x 1.5 x 3	0.9 x 0.9 x 4
Slice orientation	Coronal, transverse, sagittal	Coronal, transverse, sagittal
No. slices	36	25
Fast imaging mode	TSE (SPLICE)(factor 69)	EPI (factor 63)
SENSE	yes, factor = 1.8	yes, factor = 2
Slice gap (mm)	0	1
TE (ms)	'shortest'	'shortest'
TR (ms)	'shortest'	10000
Flip angle (degrees)	90	90
Total scan duration (min)	10:01	2:00
Act. TR/TE (ms)	8125 / 68	10000 / 101
Reference tissue	Grey matter	Liver
b-values	0, 100, 200, 800	0, 500, 900, 2000

Table 3.2: Most relevant parameters of the two DWI protocols: LUMC-SPLICE and QIBA-EPI.

3.2. mDIXON Quant Phantom Measurements

3.2.1. Calimetrix Phantom

The Calimetrix PDFF phantom (Figure 3.2) is a commercially available spherical phantom designed for testing and quality control of fat quantification (Fat Fraction Phantom, Model 300; Calimetrix [34]). It consists of twelve 22.5 mL volume agar gel-based vials for which the nominal true FF values are selected to represent the relevant biologic range in human liver (0% to ~50%). One of the vials houses a substance with a FF of exactly 100%. The MR-verified values of the used phantom are given in Table 3.3. Motion-related artifacts and air bubbles - causing susceptibility artifacts - are avoided because of its gel state.

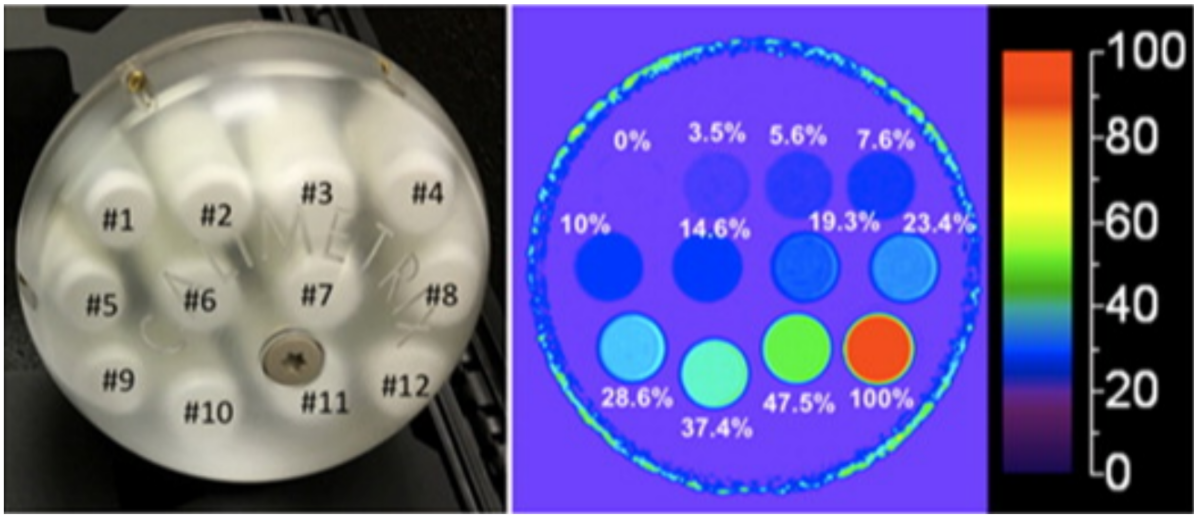


Figure 3.2: The Model 300 Calimetrix PDFF phantom [29]. The FF values given in the right image are different from the values in the phantom we used. Each product has original FF values.

Vial position	1	2	3	4	5	6	7	8	9	10	11	12
Fat Fraction (%)	0	2.7	5.3	7.8	10.1	15.5	20.4	23.6	30.2	39.9	50.1	100

Table 3.3: The MR-verified fat fractions within a tolerance of 1.5% (absolute FF) for the Calimetrix phantom that is used. Vial 1 contains no triglycerides and the FF value is exact. Vial 12 contains pure oil and the FF value is exact.

3.2.2. MR Image Acquisition

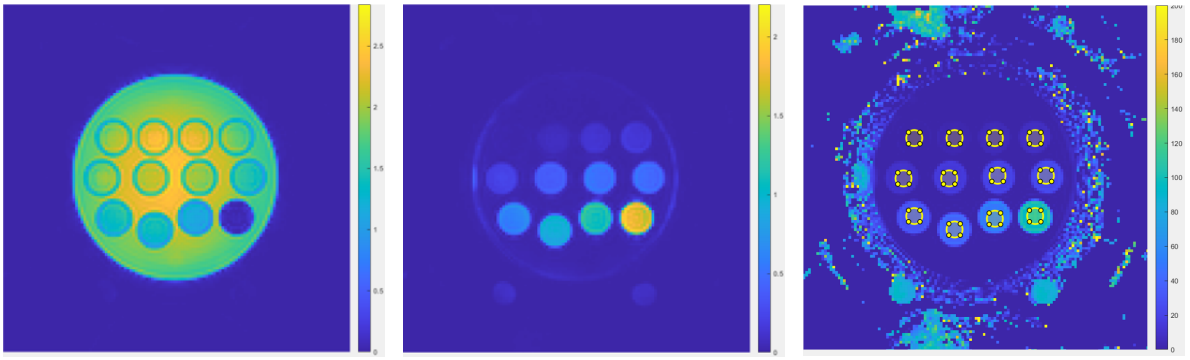
Two imaging experiments were done with the Calimetrix PDFF phantom on the same Philips Ingenia 3T scanner at LUMC. The phantom was scanned on two separate days, two weeks apart. For both protocols, the LUMC muscle and the standard Philips protocol (Table 3.1), three-dimensional volumetric MRI data sets were acquired in transverse direction. During all measurements, the phantom was positioned in a way that the vials were aligned with the main magnetic field, the z-axis. Image acquisition was done in the axial direction, capturing cross sections of the twelve vials. A Philips 16-channel head coil was used. The phantom was placed in the MRI room one hour prior to scanning to adjust to the room temperature.

Additionally, measurements were done with longer repetition times, since we want to investigate the influence of this parameter. Since SNR typically increases and T1 effect typically decreased, we expect a better approximation of the true fat-to-water ratio. In both protocols, solely the TR was manually adjusted; for the LUMC muscle protocol from 12 ms to 36 ms and for the standard Philips protocol from 5.6 ms to 17 ms.

3.2.3. ROI Segmentation and Analysis

Quantitative parametric FF maps were reconstructed by Philips' online software. We used the Philips reconstruction on for all the measurements. However, to compare the performance of both reconstruction methods, we additionally applied the in-house LUMC reconstruction method on raw acquisition data of one scan by the LUMC muscle protocol. The resulting FF values are compared to the reconstruction outcomes of the standard Philips method. This was done only for one slice in one scan since the reconstruction is quite computationally expensive.

For segmentation, circular regions of interest with a diameter of around 1 cm were drawn on the FF maps to segment vial areas in the cross section images. The radius of the circular segmentations was carefully chosen; large enough to cover many voxels, but not too large so that the edge was not touched in any of the slices. Typically, four slices (of each 3 mm) in the middle of the vials were used for segmentation. The segmentation was manually done with the use of a Matlab script. Output images of the mDIXON Quant and segmentation circles can be seen in Figure 3.3. The Matlab script was also used for calculation of the mean FF value and standard deviation for each ROI in the Philips FF map.



(a) Water image. (b) Fat image. (c) Philips FF map and segmentation.

Figure 3.3: (a), (b) and (c) are output images from the mDIXON Quant sequence. In (c), the segmentation of the vials can be seen. Colorbars are on the right for each image.

3.3. mDIXON Quant Healthy Volunteer Testing

3.3.1. Subjects

Two female subjects with the age of 50 years and BMI of approximately 29 were included. The subjects did not have a history of radiotherapy in the head and neck region and did not have any (suspected) disease of the salivary glands or swallowing muscles. Therefore, we assume that these organs can be considered as healthy tissue. There were no MR contradictions.

3.3.2. MR Image Acquisition

Data acquisition was done using two 3.0 T MR scanners (Ingenia, Philips Medical Systems, Best, The Netherlands) on two seperate locations. The best performing protocol from the phantom experiment, which was the standard Philips protocol, was used in this study. An overview of the main MRI parameters is shown in table 3.1. Volunteer 1 (V1) and volunteer 2 (V2) were scanned in Leiden University Medical Centre (LUMC) with a 16-channel head-and-neck coil. Two weeks later, only V1 was scanned again in AVL (Antoni van Leeuwenhoek Hospital, Amsterdam) with a 16-channel head-and neck coil. These measurements in AVL were repeated twice with only V1. In one of those scanning sessions, the protocol was performed twice in a row, to verify short-term repeatability. There were no MRS measurements done as reference standard.

Halfway the first two scanning sessions of V1, an acid citric drink was given to the volunteer in order to stimulate the parotid glands. After stimulation, the parotid glands were expected to produce more saliva than they do at rest and thus show a decrease of the FF value. Saliva production of the other glands is more constant, so stimulation was not expected to cause any FF differences. The FF values of the other tissues of interest were also not expected to change significantly. The protocol was used once prior to the stimulation and once after the stimulation in both scanning sessions. The stimulation included repositioning of the volunteer, since the table was required to slide out and in and the volunteer was required to sit up to drink. The acid citric stimulation was not done for V2. An overview of all the measurements can be seen in Figure 3.4.

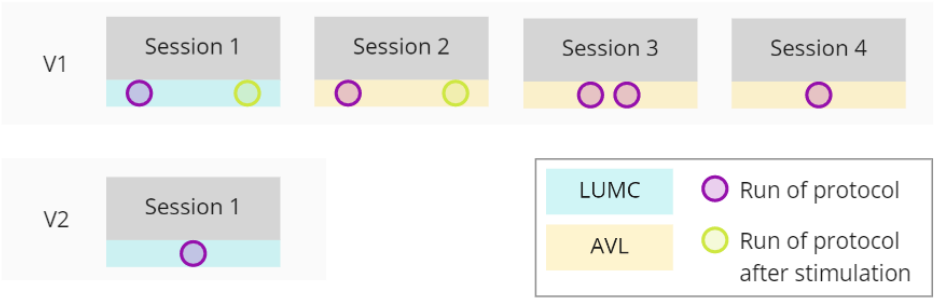


Figure 3.4: Overview of the scanning sessions done with V1 (volunteer 1) and V2 (volunteer 2).

3.3.3. ROI Segmentation and Analysis

The FF maps of the mDIXON Quant are automatically calculated and given as output by the Philips software. All FF maps were loaded in Raystation 10B Research (RaySearch; Laboratories AB Stockholm, Sweden) for image analysis. Delineations of the organs of interest were made on T1 Dixon scans according to the LIPPv2.1 guidelines. The organs that were delineated were the parotid, submandibular and sublingual glands of both sides, the oral cavity and the pharyngeal constrictor muscles (PCM) (see Figure 3.5). Additionally, part of the masseter muscle and a region of fat in the cheeks was segmented on both sides.

Subsequently, the resulting delineations were copied to the mDIXON Quant FF maps where they were manually reduced along the edges of the organs in order to rule out partial volume effects and to be more certain that the segmented volume contains mostly pixels of the tissue of interest. Furthermore, the external carotid artery was excluded from the parotid gland segmentations, since these voxels in low-fat tissue would decrease the mean FF value (see Figure 3.6).

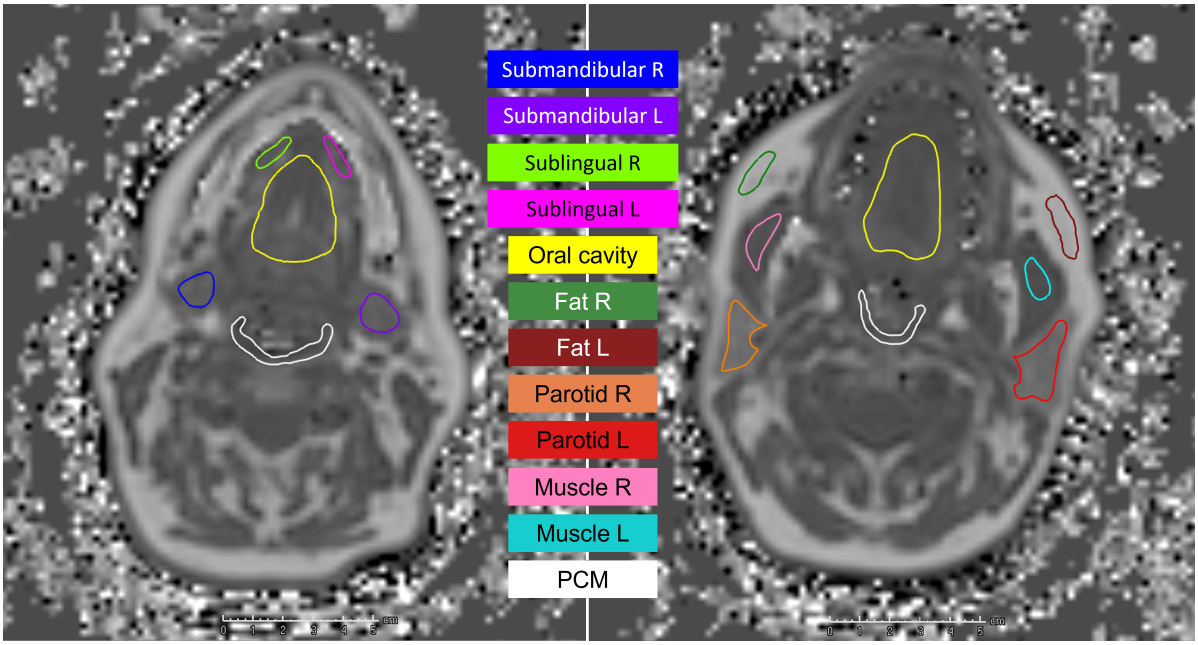


Figure 3.5: Example of delineations in one volunteer in two axial slices. **Left:** Axial slice near the chin. **Right:** Axial slice near the mouth.

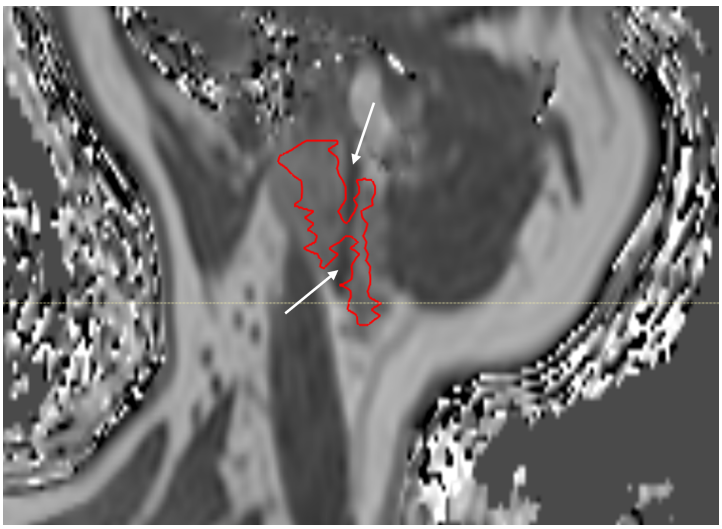


Figure 3.6: Sagittal view of right parotid gland in volunteer 1 with delineation. In this view, it can clearly be seen that the external carotid artery (indicated with the white arrows) is excluded from the delineation.

The mean and standard deviation of the FFs in these ROIs are calculated using Matlab script in the Raystation software. The coefficient of variation (CV) is calculated based on the mean FF values and standard deviations for all tissues of interest and is used as a tool to assess variation in the repeated measurements. This coefficient, also known as relative standard deviation (RSD), is defined as the ratio of the standard deviation to the mean and can be used as a standard measure for dispersion. Acceptable levels of CV depend on the research sector. Generally, a CV lower than 10% is considered to represent a relatively low spread of data values relative to the mean and a CV higher than 30% is considered unacceptable. We consider a CV lower than 10% to be good.

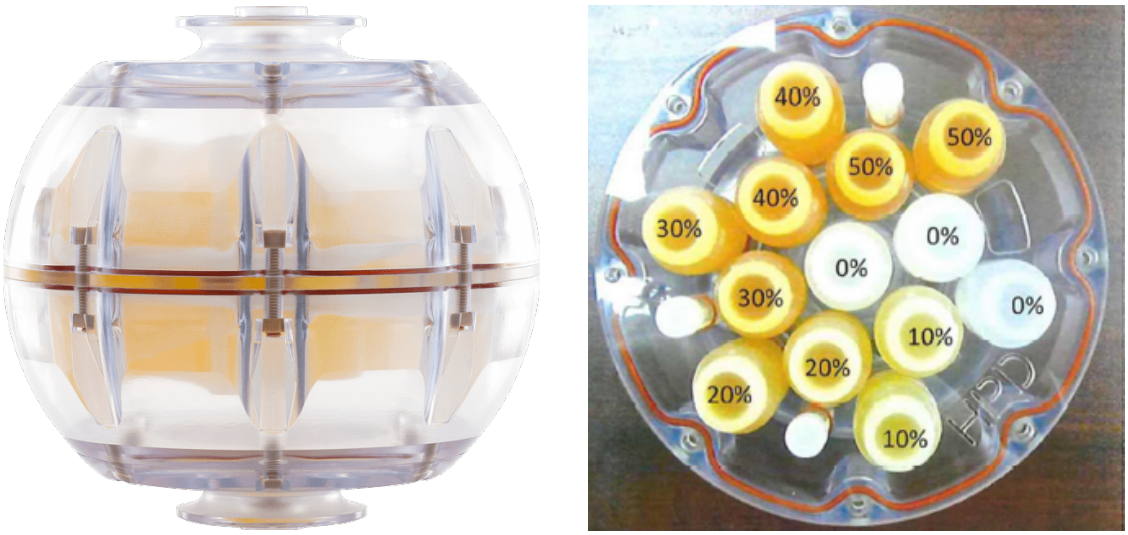
3.4. DWI Phantom Measurements

The DWI phantom experiment is done to find out whether the accuracy of the DWI-SPLICE protocol is comparable to the accuracy of the QIBA protocol on the same scanner. Furthermore, we wanted to verify whether DWI-SPLICE caused less geometrical deformations than the QIBA-EPI protocol.

3.4.1. QIBA PVP DWI Phantom

The QIBA-DWI phantom is used for this experiment. This is an ice-water diffusion phantom with an array of polyvinylpyrrolidone (PVP) solutions (Figure 3.7). The phantom was borrowed from the Netherlands Cancer Institute. Preparation of the phantom includes filling it with ice cubes and water hours prior to scanning, in order to bring the temperature of the phantom vials to 0°C. This is necessary to achieve reproducible measurements.

The 30 mL vials contain aqueous solutions with various concentrations of the polymer PVP (0, 10, 20, 30, 40 and 50%) increased PVP concentration leads to decreased ADC values. The center vial holds deionized water and should be positioned at iso-center when imaging.



(a) (b)

Figure 3.7: The QIBA DWI phantom. (a) Side view on the phantom, (b) top view on the bottom half of the phantom and the 13 vials.

3.4.2. MR Image Acquisition

Measurements were done using a 3.0 T MR scanner (Ingenia, Philips Medical Systems, Best, The Netherlands) with a 16-channel head coil. The ice-filled phantom was scanned in all three directions using two protocols; the standard QIBA protocol and the DWI-SPLICE protocol. These protocols are described in Table 3.2. During imaging, the temperature was measured inside the phantom. This temperature was not allowed to deviate more than a few tenths degrees from 0°C.

3.4.3. ROI Segmentation and Analysis

ROI segmentation was done using a software package from QIBA, which can be found in the Quantitative Imaging Data Warehouse (QIDW). The same software automatically calculates the mean ADC value per ROI. The software is deployed as a C++ compiled GUI application for 64-bit Windows systems. After segmentation, the ROI statistics are saved in a csv file, which can be opened and observed in Microsoft Excel.

4

Results

In this chapter, the findings of the different experiments are presented. In the first Section, we describe outcomes of the mDIXON Quant experiments and the performance of both protocols based on phantom experiments. The influence of a longer repetition time and different reconstruction technique is also described. Furthermore, the outcomes of healthy volunteer MR imaging are presented and compared to values found in literature. Results on the house-built phantom are not included, but can be found in Appendix C. The results of the DWI phantom experiment are given in Section 4.2.

4.1. mDIXON Quant Experiment Results

4.1.1. Calimetrix Phantom Experiment

The results from the experiment that is described in Section 3.2 are given here. An overview of the experiment outcomes can be seen in Figure 4.1. An overview of the mean FF values and standard deviation of all measurements done can be found in Appendix A.

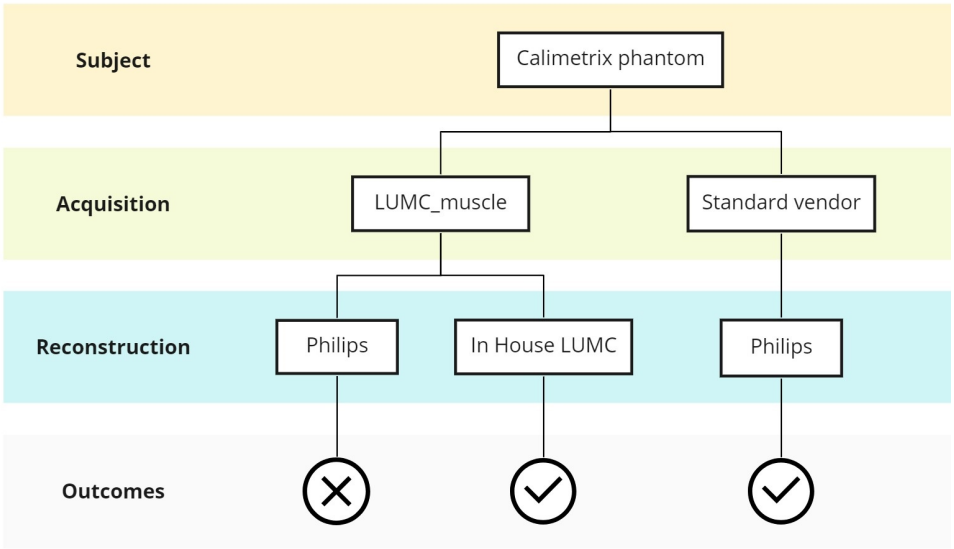


Figure 4.1: Overview of the experiments done for mDIXON Quant on the Calimetrix phantom and their outcomes. A checkmark means that accuracy and repeatability were both well. A cross means that for at least one of those, the outcomes were not good enough.

The measurements of the Calimetrix phantom are depicted in Figure 4.2. The variation between measurements of the LUMC muscle protocol was larger than the variation of the standard protocol measurements. The deviation from the true FF value was also larger for the LUMC muscle protocol measurements than for those from the standard protocol; the FF measurements were always underestimated compared to the ground truth. The deviation for the LUMC muscle protocols was increasing with the true FF value. At a true FF value of 50.1%, the deviation of the measurement was around 15%.

Thus, the accuracy of the standard protocol measurements was higher than for the LUMC muscle protocol. The variability in repetitions was quite low for both protocols. However, the standard protocol yielded results with lower variability.

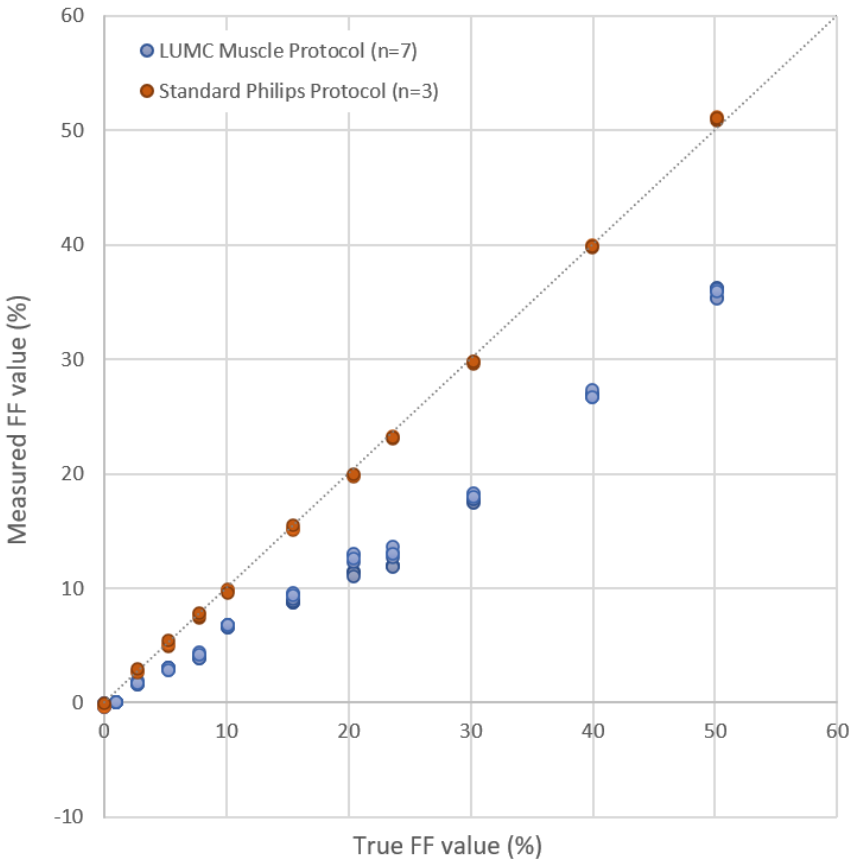


Figure 4.2: Measurements of FF values against ground truth for the two mDIXON Quant protocols. The dashed line represents a reference line on which measured value and true value are exactly the same. The dots represent each a mean FF value measured in one scan using one of the protocols; red is for the standard Philips protocol and blue is for the LUMC muscle protocol. Measurements of 100% vial are excluded in this graph for a better overview.

4.1.1.1. Effect of Longer Repetition Time

Figure 4.3 shows the measurements done by both protocols with elongated repetition time. For the LUMC muscle protocol the TR was changed from 5.6 ms to 17 ms and for the standard Philips protocol, the TR was changed from 12 ms to 36 ms, while TE1 and echo spacing remained unchanged. The deviation from the true value and the variation between measurements were in the same order as for the original protocols.

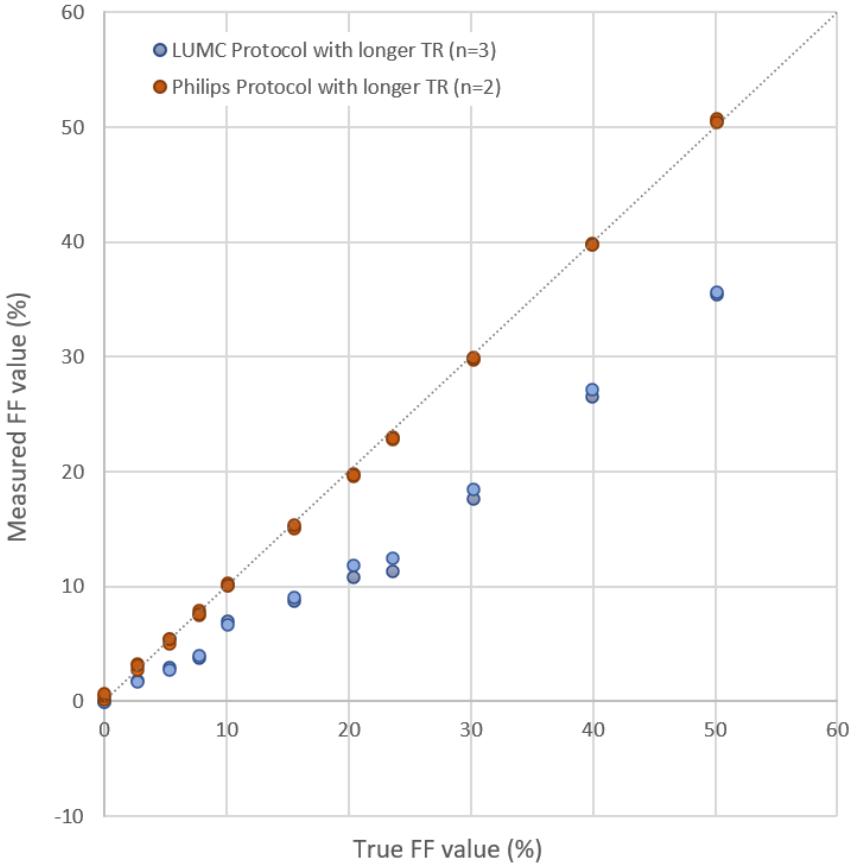


Figure 4.3: Measurements of FF values against ground truth for an adaptation of both protocols (longer repetition time). The dashed line represents a reference line on which measured value and true value are exactly the same. The dots represent each a mean FF value measured in one scan using one of the protocols; red is for the Philips protocol and blue is for the LUMC protocol. Measurements of 100% vial are excluded in this graph for a better overview.

4.1.1.2. Effect of alternative reconstruction

Figure 4.4 shows the results of the two different reconstruction methods in combination with the acquired data of one image. Even though acquisition was done with the same method - the LUMC muscle protocol, the in-house reconstruction method from LUMC produced better results for all vials.

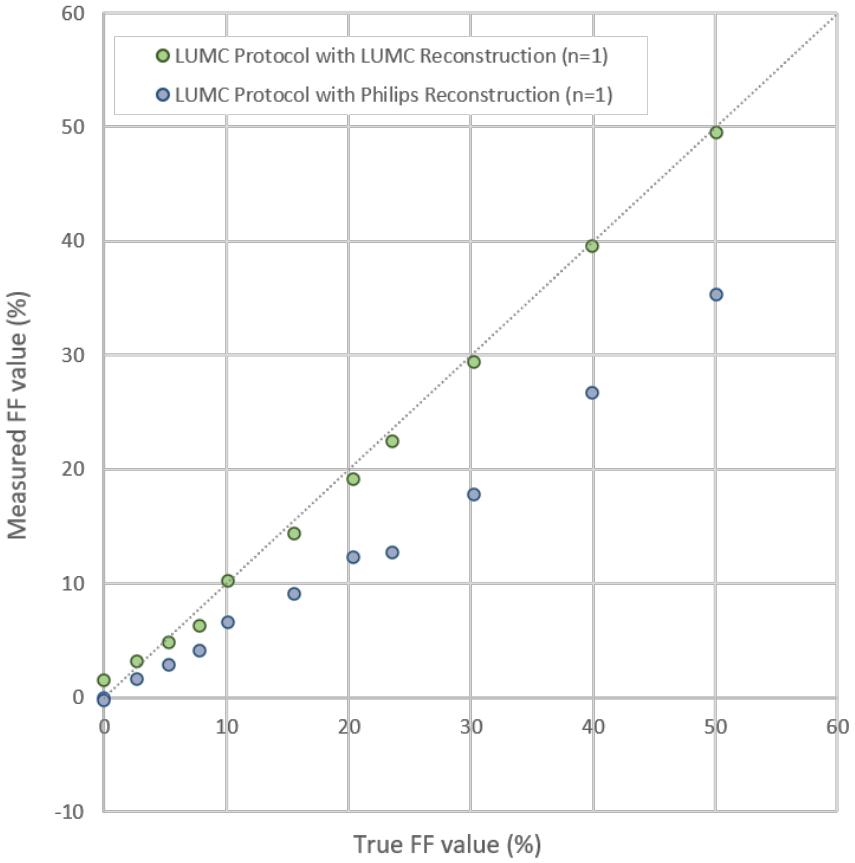


Figure 4.4: Measurements of FF values against ground truth for the LUMC protocol with Philips reconstruction and LUMC reconstruction. The dashed line represents a reference line on which measured value and true value are exactly the same. The dots represent each a mean FF value measured in one scan using one of the protocols; green is for the LUMC reconstruction and blue is for the Philips reconstruction. Measurements of 100% vial are excluded in this graph for a better overview.

4.1.2. Healthy Tissue Measurements

Here, the results for the experiment of Section 3.3 are given. An overview of the mean FF values and standard deviations for all measurements done can be found in Appendix A. Based on the outcomes of the phantom measurements, we only used the standard protocol for this analysis. The results of the first session in LUMC are given in Figure 4.5. In this session, V1 and V2 were scanned. For all tissues of interest, the measured FF value was higher for volunteer 2.

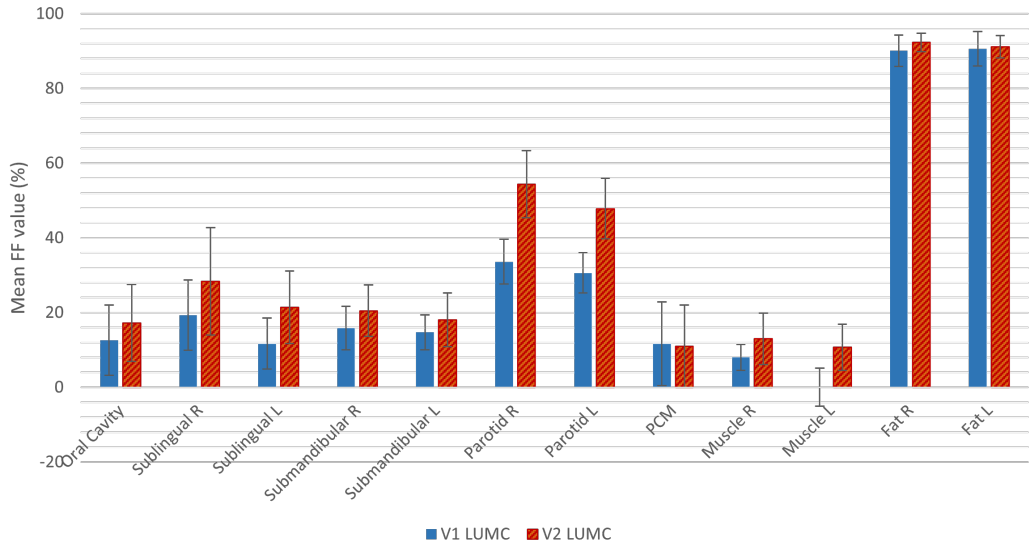


Figure 4.5: FF measurement results for the LUMC scanning session for V1 (before stimulation) and V2.

4.1.2.1. Salivary gland stimulation

The mean FF value and standard deviations of the first two scanning sessions including acid citric stimulation with volunteer 1 can be seen in Figure 4.6. No clear changes in FF values were found for the parotid glands before and after stimulation.

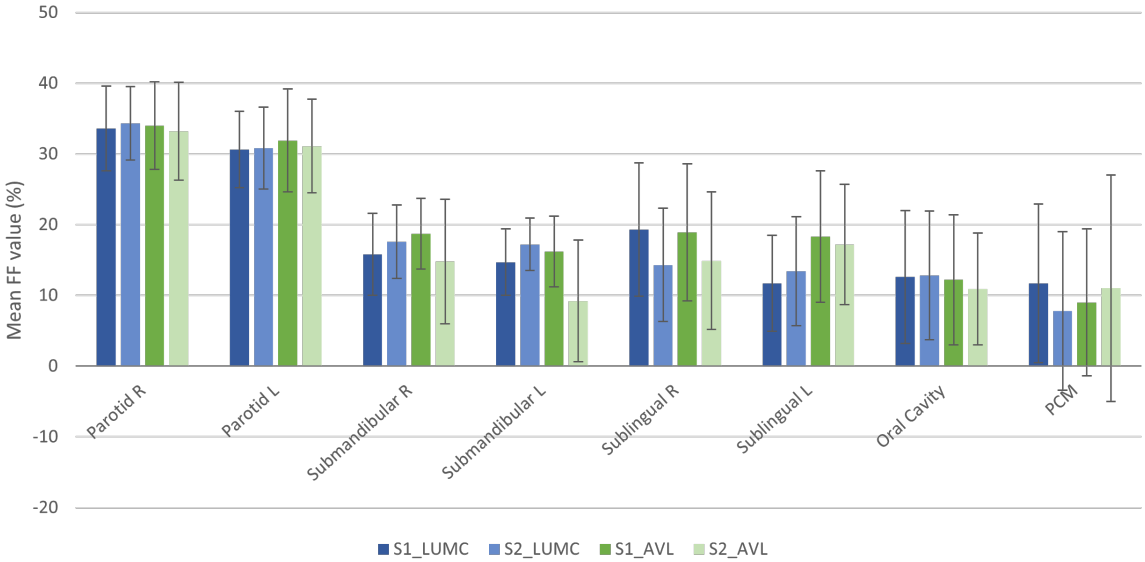


Figure 4.6: Mean FF value (%) of tissues of interest for the standard protocol (volunteer 1 only). The error bars show the standard deviation. Measurements have been done in LUMC and AVL before and after acid citric stimulation (S1 is before stimulation and S2 is after stimulation).

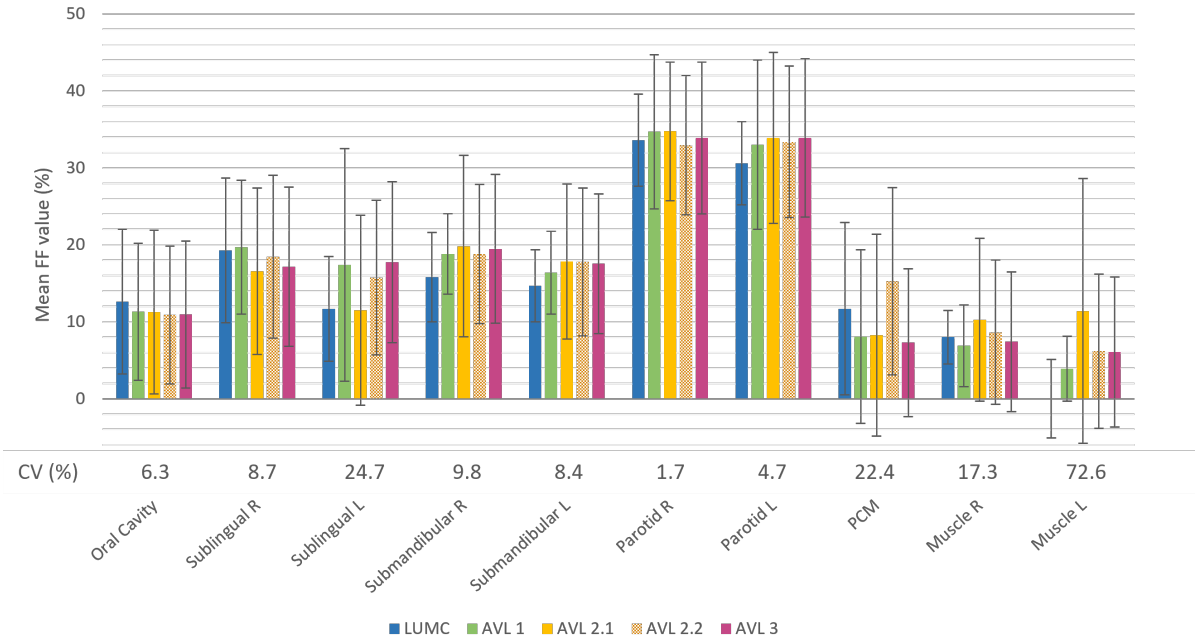


Figure 4.7: Mean FF value (%) of tissues of interest for the standard protocol of measurements done in LUMC and AVL on Volunteer 1. Standard deviations are depicted by the error bars. The CV (coefficient of variance) is given per tissue of interest.

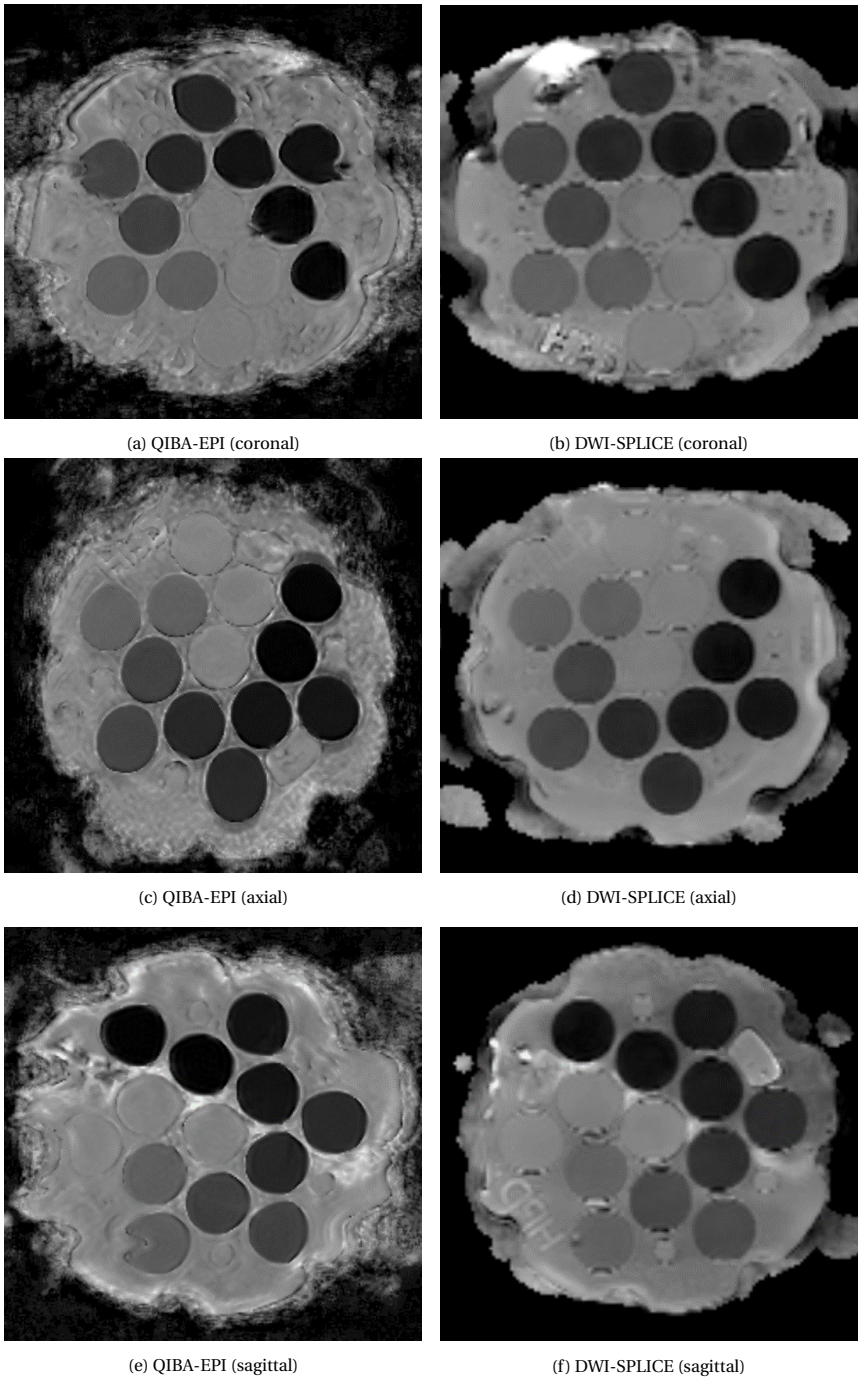
4.1.2.2. Repeatability

In total, the standard Philips protocol has been performed on Volunteer 1 on four different days, resulting in a total of five scans (on one day, the protocol was scanned twice). The resulting mean FF values for the tissues of interest are shown in Figure 4.7. The coefficient of variation (CV) was measured for all tissues of interest. Except for the left sublingual gland, the PCM and the muscle segmentations on both sides, all CV's were below 10%.

4.2. DWI-SPLICE Performance

4.2.1. Geometrical Accuracy

The geometrical accuracy of the vials in the DWI-SPLICE images was better than that of the vials in the QIBA-EPI images for all directions. See Figure 4.8 for the difference of geometrical accuracy in the center slice of all three directions.



(a) QIBA-EPI (coronal) (b) DWI-SPLICE (coronal)

(c) QIBA-EPI (axial) (d) DWI-SPLICE (axial)

(e) QIBA-EPI (sagittal) (f) DWI-SPLICE (sagittal)

Figure 4.8: A center slice of the ADC maps from both protocols in all three directions. Window levels are the same in all images.

4.2.2. ADC values

In Table 4.1, the deviation of each measurement to the ground truth is presented. The deviation of the measured ADC values is:

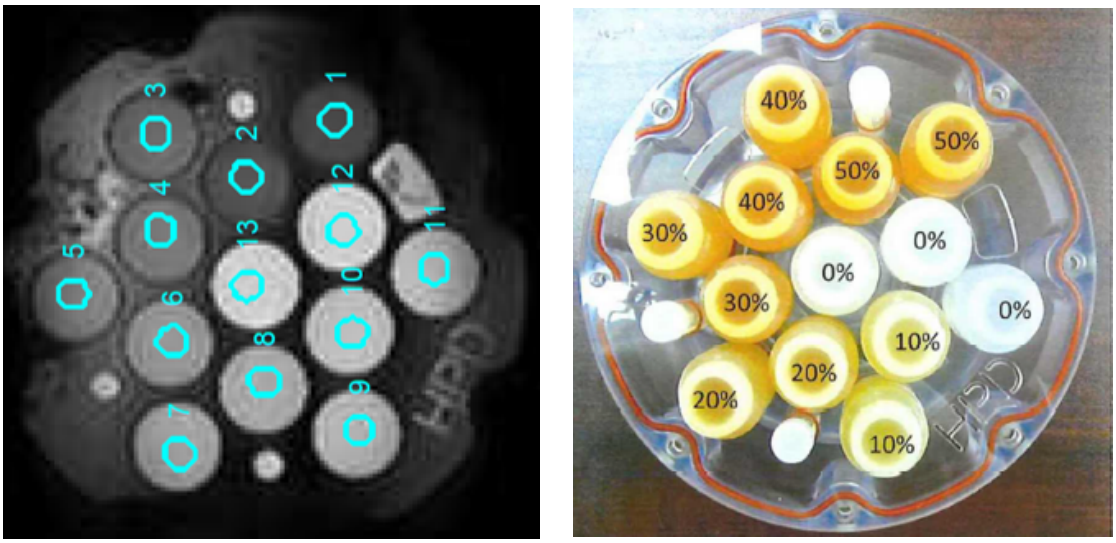
$$Deviation(\%) = \frac{ADC_{Measured} - ADC_{True}}{ADC_{True}} \cdot 100\% \tag{4.1}$$

The deviation was often higher than 5% for ADC values lower than 0.8 x 10^{-3} mm^2/s. For higher ADC values, both sequences yielded comparable results.

			Deviation (%)					
		ADC	QIBA-EPI			DWI-SPLICE		
vial	% PVP	($x10^{-3}$ mm^2/s)	COR	AX	SAG	COR	AX	SAG
1	50 (o)	0.125	19.8	10.0	4.2	88.0	35.0	30.6
2	50 (i)	0.129	24.2	5.4	7.6	90.0	28.9	82.1
3	40 (o)	0.244	6.0	3.3	3.6	0.4	1.8	14.3
4	40 (i)	0.239	1.2	3.4	2.6	4.3	8.3	33.7
5	30 (o)	0.402	4.6	1.9	3.2	13.3	3.0	13.3
6	30 (i)	0.402	1.9	0.1	2.4	1.2	2.0	14.0
7	20 (o)	0.599	1.3	2.2	2.2	6.0	3.5	8.6
8	20 (i)	0.601	0.9	0.3	2.0	3.3	2.2	5.4
9	10 (o)	0.833	1.6	3.2	2.6	2.4	2.3	0.2
10	10 (i)	0.840	2.1	0.3	1.1	2.3	0.5	0.5
11	0 (o)	1.118	4.8	1.2	1.4	2.0	0.9	0.7
12	0 (i)	1.105	2.1	2.3	1.2	3.8	4.5	0.8
13	0 (c)	1.109	2.4	0.9	1.2	4.1	5.1	0.9

Table 4.1: Deviation (%) from true ADC value (x 10^{-3} mm^2/s) for 6 scans. The first 3 are from the QIBA-DWI sequence and the final 3 from the DWI-SPLICE sequence. Green background = lower than 3% deviation, Orange background = between 3 and 5 % deviation and Red background = more than 5% deviation. (i) is inner vial, (o) is outer vial and (c) is center vial.

An example of drawn ROIs can be seen in Figure 4.9a. Since the anatomy of the DWI-SPLICE images was much better preserved, the segmentation was done easier for the DWI-SPLICE images.



(a) Regions of interest on a DWI-SPLICE acquisition image.

(b) Top view on phantom vials with percentage of PVP for each vial.

Figure 4.9: Overview of vial numbers and corresponding PVP percentage.

5

Discussion

Fat fraction (FF) and apparent diffusion coefficient (ADC) values estimated by mDIXON Quant and DWI-SPLICE techniques respectively, are relatively new quantitative imaging parameters and increasingly accepted as imaging biomakers for all sorts of purposes. The BOCASEcA study aims to research whether these techniques can be used as biomarkers for patient-reported xerostomia and dysphagia post-radiotherapy. In this project, steps have been taken to validate the use of certain protocols in the BOCASEcA study. For mDIXON Quant, we performed phantom studies and a healthy volunteer study to evaluate the accuracy and repeatability of the Philips technique by using the standard default protocol and a protocol that is used in LUMC on muscles throughout the whole body. Another phantom study was done for DWI-SPLICE performance evaluation. This technique is known to have less susceptibility issues than conventional EPI-DWI techniques. We tested both the accuracy and deforming artifacts for a QIBA-EPI sequence and a clinically used DWI-SPLICE protocol from LUMC.

This chapter reflects on the outcomes of the research. In the first three Sections, outcomes are summed up, conclusions are drawn, possible explanations are suggested and limitations are discussed for each experiment. Section 5.4 proposes recommendations for further research.

5.1. mDIXON Quant Phantom Measurements

The standard mDIXON Quant protocol from Philips can be considered as an accurate and robust tool for fat quantification. The LUMC muscle protocol however, performed less well using the Philips reconstruction. The underestimation of this latter protocol was proportional with the true FF value. A higher variation within multiple measurements was observed for the LUMC muscle protocol than for the standard Philips protocol. Tripling the repetition time did not influence the performance of the protocols.

The results are in line with a study done by Hu et al. [29]. The authors have successfully examined the linearity and range of bias in the Model 300 Calimetrix phantom at nine participating centers. They found that Philips systems that use the default vendor protocol perform within their defined acceptance bias margins both at 1.5 T and 3.0 T. However, the slope of a custom created protocol (with longer echo times) was biased from the identity line. Their measured FF values were an underestimation compared to true FF values as well. The deviation was also proportional with the true FF value.

The use of an off-scanner post-processing Matlab script, created by K. Keene and A. Sardjoe Mishre for LUMC, combined with the LUMC muscle protocol led to better outcomes. Philips' mDIXON Quant acquisition is a vendor product and their post-processing script is not open source. Therefore, it is hard to exactly determine what computational steps and corrections are included in their pipeline. For this reason, the causes of the difference in outcomes by the two protocols cannot be easily found. However, some confounders could potentially explain the differences and there are a few that can be ruled out.

First of all, incorrect B0 estimation cannot explain the FF calculation differences of this magnitude. However it could lead to more extreme differences because it can cause fat-water swaps. To counter these fat-water swaps, Philips is additionally using a nearest neighbour or region-growing technique in their water-fat separation method, which works relatively well. Another effect that cannot explain the differences is a bias caused by regional B1 differences. Fat and water differ in T1 times, thus T1 can influence the FF calculation, dependent on the flip angle and repetition time. However, the FF differences caused by T1 weighting are modelled to be maximally 2% given the mentioned acquisition parameters and are not dependent on echo times. Lastly, the signal-to-noise ratio (SNR) is presumably also not the cause of the FF calculation differences, since it is almost independent on echo time as well.

Since the most notable differences in the two protocols are the first echo time and the echo spacing, we focused on factors that are modulated by these acquisition parameters. Two factors that are highly echo time dependent are T2* and eddy current effects. Firstly, the Philips mDIXON Quant includes T2* in the water-fat separation method and creates a T2* map as an output. The mDIXON Quant corrects for the influence of T* by including it in the signal model. Thus it is unlikely that the differences in FF can be assigned to T2* effects.

Eddy current effects also increase when shorter echo times are used. The Philips reconstruction compensates for these effects [28], but no specifications of this correction algorithm could be found in literature. If this correction is optimized for short echo times and unable to perform well on protocols with longer echo times, it could explain the difference in performance between the two protocols in this experiment.

There are a few limitations for this experiment. Firstly, the LUMC muscle protocol was repeated seven times and the standard Philips protocol was repeated only three times. More repetitions of the standard Philips protocol would be desired in order to make a stronger statement and to paint a more accurate picture of the spread. Furthermore, the measurements were all done on only one scanner. However, for assessing the robustness of the technique regardless of the scanning system, it should be tested on different scanners. Lastly, the Calimetrix phantom is supposed to be scanned at room temperature (according to its accompanying manual). However, the temperature inside the phantom could not be measured before, during and after scanning. Increasing temperature due to scanning could have had an effect on the outcomes, since the proton resonance frequency of water relative to fat is temperature dependent [35]. For this reason, a change in temperature could lead to inaccuracies in spectral fat modeling.

5.2. mDIXON Quant Healthy Volunteer Testing

The results of head-and-neck FF measurements using the standard Philips protocol were promising. The standard mDIXON Quant protocol produced repeatable measurements for most organs, but for the smallest organs, the coefficient of variation was higher than 10% and thus robustness was considered to be insufficient. Stimulation by an acid drink did not lead to clear consistent differences in FF measurements.

The sublingual glands and PCM were difficult to delineate. These organs are small and complicated to distinguish from nearby tissues. This could lead to bigger delineation errors and more variation in the measurements. Therefore it is not unexpected that for exactly these organs, the CV was quite high. The CV for the muscle tissue on both sides was high as well. This could be explained by the relatively low FF values that were measured in a relatively small segmentation. It is unexpected however, that the variation in muscle was this large, as the volume appeared quite homogeneous on the FF map.

Stimulation probably has no to small influence on the FF value of the parotid glands as of any other tissue of interest. However, the measurements were done a few minutes after stimulation; around 10 minutes in the first session in LUMC and about 30 minutes in the second session, in AVL. Therefore, the effect of the stimulation could already be reduced to a minimum by the time the sequences were run.

Based on the measurements for volunteer 1, the average FF values of the parotid glands were comparable to numbers found in literature. For the submandibular glands of V1, the FF values that we found were about 10% higher than what was found in literature. The mean FF value for the oral cavity of V1 was for all five measurements just a bit higher than 10%. This is lower than the FF value that Humbert et al. found for the tongue, which was 26.5% [20].

Short-term repeatability could have been tested more strongly, by measuring twice in a row more often. Furthermore, the delineations for the small organs are uncertain, since they are difficult to delineate.

5.3. DWI Phantom Measurements

The DWI-SPLICE protocol yielded better geometrical accuracy than the QIBA-EPI protocol, as expected. However, accuracy of the DWI-SPLICE measurements was lower in general than for the QIBA-EPI protocol. For both protocols, the deviation from the true ADC value was highest for the vials with the lowest ADC value, thus the values with the least water mobility. The DWI-SPLICE sequence performed worse for those vials with low ADC values compared to the QIBA-EPI protocol. However, the accuracy was acceptable for ADC values higher than approximately 0.6 x 10^{-3} mm^2/s, which is the biologically relevant range in head-and-neck tissues. Therefore, the DWI-SPLICE protocol is expected to perform reasonably well for the purpose of qMRI in the head-and-neck area, based on these simple phantom measurements.

The larger deviation from the true values for the DWI-SPLICE protocol is not unexpected, since we expect a lower SNR compared to EPI-DWI techniques. However, we did not perform any SNR measurements to prove that SNR was actually lower. Furthermore, we had expected to find that the outer vial measurements would yield larger deviations from the true ADC values than the inner vial, but this pattern cannot be recognized for our measurements. Incorrect preparation and imaging of the phantom could lead to incorrect measurements. For example, the ice should keep the inside of the phantom at 0°C during the whole imaging time. Furthermore, since DWI is typically susceptible to B0 homogeneities, phantom centering is crucial. Even though preparation and imaging were done carefully, it could be that slight temperature changes or errors in positioning have led to deviations larger than the system errors.

The test-retest repeatability was not assessed, since for all directions and protocols only one scan was made. In order to find out whether the outcomes are not coincidental, repeating the measurements is necessary. In addition, the protocols were not tested in-vivo. This step is required to find the best parameters for the DWI-SPLICE protocol that will lead to accurate and robust ADC mapping in healthy volunteers and patients. One of the downsides of our DWI-SPLICE protocol however, is that it is relatively long for patient imaging.

5.4. Recommendations

The standard Philips mDIXON Quant protocol is accurate and robust and can be used for the BOCASEcA study. Some parameters can be adjusted if necessary, e.g. resolution or FOV, but this should be done with caution. For example, if the resolution is improved (i.e. voxel size is reduced), this could lead to an increase of the shortest possible echo times. Increased echo times could lead to worse accuracy. Therefore, if any desired adjustments of the protocol have influence on the echo times, it is recommended to perform the protocol on the Calimetrix phantom first to test the accuracy. The total scan time does not need to be taken into account if adjustments are made, since the sequence is already only half a minute long and breath hold - the cause for the brevity of the sequence - is unnecessary for head and neck imaging.

Moreover, since multiple scanners will be used in the BOCASEcA study, it would be valuable to test the performance of the scanner and protocol combination for each scanner on the Calimetrix phantom.

For the DWI-SPLICE, the outcomes seem promising. However, it is recommended to validate robustness on a phantom and to optimize the protocol for clinical use by testing it on healthy volunteers. Our DWI-SPLICE protocol is relatively long for in-vivo experiments, so it is recommended to shorten the scan time by e.g. lowering the NSA for various b-values or degrading the resolution.

In summary, through the phantom and volunteer experiments that were carried out, the performance of two mDIXON Quant protocols was inspected. Adequate accuracy and robustness was found for the standard Philips protocol. Furthermore, the accuracy of a DWI-SPLICE protocol was validated through a phantom designed by the Quantitative Imaging Biomarkers Alliance. Since their accuracy was validated, it is recommended that both protocols are further optimized for use in the BOCASEcA study.

A

Appendix A: Measurements

A.1. Calimetrix Measurements

LUMC Muscle

		TR / TE1 / ΔTE (ms) = 12 / 2.3 / 1.5								
		June 24th			June 7th					
		1	2	3	4	5	6	7		
	TRUE	mean	mean	mean	mean	mean	mean	mean	Average	Bias (%)
1	0.0	-0.1	-0.3	-0.3	-0.2	0.0	-0.2	-0.3	-0.2	0.2
2	2.7	1.7	1.6	1.7	1.6	1.9	1.9	1.7	1.7	1.0
3	5.3	3.0	3.0	3.1	2.8	3.0	3.1	2.9	3.0	2.3
4	7.8	3.9	4.0	4.2	4.1	3.9	4.4	4.2	4.1	3.7
5	10.1	6.7	6.6	6.8	6.6	6.8	6.8	6.7	6.7	3.4
6	15.5	8.7	8.8	8.9	9.1	9.5	9.6	9.4	9.1	6.4
7	20.4	11.4	11.3	11.1	12.3	12.9	13.0	12.6	12.1	8.3
8	23.6	12.0	12.0	11.8	12.7	13.1	13.6	13.0	12.6	11.0
9	30.2	17.6	17.5	17.8	17.8	18.2	18.3	18.0	17.9	12.3
10	39.9	26.9	26.8	26.8	26.7	27.0	27.3	26.7	26.9	13.0
11	50.1	36.2	36.2	35.9	35.4	35.3	36.1	35.9	35.9	14.2
12	100.0	101.5	101.5	101.5	102.3	102.2	102.6	102.8	102.0	-2.0

LUMC Long TR

		36 / 2.3 / 1.5			
		1	2		
	TRUE	mean	mean	Average	Bias (%)
1	0.0	0.0	0.0	0.0	0.0
2	2.7	1.8	1.7	1.7	1.0
3	5.3	3.0	2.8	2.9	2.4
4	7.8	3.7	3.9	3.8	4.0
5	10.1	6.9	6.7	6.8	3.3
6	15.5	8.8	9.0	8.9	6.6
7	20.4	10.8	11.8	11.3	9.1
8	23.6	11.4	12.4	11.9	11.7
9	30.2	17.6	18.4	18.0	12.2
10	39.9	26.5	27.1	26.8	13.1
11	50.1	35.4	35.7	35.5	14.6
12	100.0	101.5	101.7	101.6	-1.6

Standard Philips

		5.6 / 0.97 / 0.7				
		June 24th				
		1	2	3		
	TRUE	mean	mean	mean	Average	Bias (%)
1	0.0	-0.2	-0.3	-0.1	-0.2	0.2
2	2.7	2.6	2.7	3.0	2.8	-0.1
3	5.3	5.0	5.1	5.5	5.2	0.1
4	7.8	7.5	7.7	7.8	7.6	0.2
5	10.1	9.9	9.8	9.6	9.8	0.3
6	15.5	15.1	15.2	15.5	15.3	0.2
7	20.4	19.7	19.8	19.9	19.8	0.6
8	23.6	23.1	23.2	23.2	23.1	0.5
9	30.2	29.6	29.8	29.8	29.7	0.5
10	39.9	39.8	39.9	39.9	39.9	0.0
11	50.1	50.9	51.2	51.0	51.0	-0.9
12	100.0	97.8	98.0	98.5	98.1	1.9

Philips Long TR

		17 / 0.97 / 0.7				
		June 24th				
		1	2	3		
	TRUE	mean	mean	mean	Average	Bias (%)
1	0.0	0.2	0.5	0.6	0.6	-0.4
2	2.7	2.8	3.2	3.2	3.2	-0.4
3	5.3	5.0	5.5	5.4	5.5	0.0
4	7.8	7.5	7.9	7.6	7.8	0.1
5	10.1	10.1	10.3	10.1	10.2	-0.1
6	15.5	15.1	15.1	15.4	15.2	0.3
7	20.4	19.6	19.8	19.7	19.7	0.7
8	23.6	22.8	23.0	22.9	23.0	0.7
9	30.2	29.8	29.8	29.9	29.9	0.4
10	39.9	39.7	39.9	39.7	39.8	0.1
11	50.1	50.5	50.7	50.4	50.5	-0.4
12	100.0	98.1	97.8	98.2	98.0	1.9

Table A.1: Calimetrix measurements (FF values in %)

A.2. Healthy volunteers

	V1_LUMC				V1_AVL 1									
	Before stimulation		After stimulation		Before stimulation		After stimulation		V1_AVL 2.1		V1_AVL 2.2		V1_AVL 3	
	mean	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**
Oral Cavity	12.6	9.4	12.8	9.1	11.3	8.9	10.9	7.9	11.3	10.7	10.9	8.9	11.0	9.5
Sublingual R	19.3	9.4	14.3	8.0	19.7	8.7	14.9	9.7	16.6	10.8	18.5	10.6	17.2	10.4
Sublingual L	11.7	6.8	13.4	7.7	17.4	15.1	17.2	8.5	11.5	12.3	15.8	10.1	17.8	10.5
Submandibular R	15.8	5.8	17.6	5.2	18.8	5.2	14.8	8.8	19.8	11.8	18.8	9.1	19.5	9.6
Submandibular L	14.7	4.7	17.2	3.7	16.4	5.4	9.2	8.6	17.8	10.1	17.8	9.6	17.5	9.1
Parotid R	33.6	6.0	34.3	5.2	34.7	10.0	33.2	6.9	34.8	9.0	32.9	9.0	33.9	9.9
Parotid L	30.6	5.4	30.8	5.8	33.0	11.0	31.1	6.6	33.9	11.1	33.4	9.8	33.9	10.3
PCM	11.7	11.2	7.8	11.2	8.1	11.3	11.0	16.0	8.3	13.1	15.3	12.2	7.3	9.6
Muscle R	8.0	3.5	7.3	3.2	6.9	5.3	3.4	5.2	10.3	10.6	8.6	9.4	7.4	9.1
Muscle L	0.0	5.1	3.4	2.8	3.9	4.2	7.6	3.8	11.4	17.2	6.2	10.0	6.1	9.8
Fat R	90.1	4.2	87.6	7.5	90.7	3.9	86.3	5.5	-	-	-	-	-	-
Fat L	90.6	4.6	90.0	5.0	90.0	4.1	90.2	4.5	-	-	-	-	-	-

	V2_LUMC	
	mean	**std**
Oral Cavity	17.2	10.3
Sublingual R	28.4	14.4
Sublingual L	21.4	9.7
Submandibular R	20.5	6.9
Submandibular L	18.1	7.2
Parotid R	54.4	9
Parotid L	47.8	8.1
PCM	11	11
Muscle R	13	6.9
Muscle L	10.7	6.2
Fat R	92.3	2.5
Fat L	91.1	3

Table A.2: Healthy volunteer measurements (FF values in %)

B

Appendix B: Fat Quantification Parameters from Literature

Table B.1 gives an overview of parameters used for multiple fat quantification studies in the head and neck region using MRI. All studies investigated FF values of salivary glands.

First author	Vendor	Type	Field strength (T)	Coil	Voxel size (mm)	Slice thickness (mm)	Flip angle (°)	FOV (mm)	No. echos	TR (ms)	TE1 (ms)	ΔTE (ms)	Total imaging time
Chang [14]	GE	Signa HDx	1.5	8-ch H&N	NR	5	10	240 x 240	3	8.9	2.0 (1)	3.5 (2), 5.0 (3)	2 min 50 s
Kise [15]	Philips	Ingenia	3.0	H&N and torso	1.5 x 1.5	2	3	230 x 230	6	10	1.3	1.0	2 min 49 s
Chikui [16]	Philips	Ingenia	3.0	H&N	1.5 x 1.5	2	3	230 x 230	6	10	1.3	1.0	4 min 15 s (SENSE)
Zhou [17]	Philips	Ingenia	3.0	16-ch H&N 32-ch torso	1.2 x 1.2	2.5	3	250 x 240	6	10	1.47	1.2	2 min 7 s
Su [18]	GE	Discovery MR750W	3.0	29-ch H&N	NR	4	5	240 x 240	6	12	4	0.66	0 min 45 s
Chu [19]	Philips	Ingenia	3.0	16-ch H&N and torso	1.1 x 1.1	2.5	3	250 x 240	6	8.7	1.44	1.2	1 min 21 s
Jimenez-Royo [36]	Siemens	Magnetom Trio	3.0	12-ch head A&P neck	NR	3	9	256 x 256	2	5.66	2.45	3.7	NR

Table B.1: Overview of relevant MR imaging parameters. Abbreviations: FOV = field of view, TR = repetition time, TE1 = first echo time, ΔTE = echo spacing, H&N = head-and-neck, A&P = anterior and posterior, NR = not reported.

C

Appendix C: Custom Made Fat-Water Phantom

Since a commercial fat-water phantom was not yet available in the beginning of this thesis project, I have built a custom made one. The construction of and MRI measurements with the phantom are elaborated on in this appendix. First, an overview is given of various construction techniques. Subsequently, the measurement experiment is explained and results are given. A conclusion is made in the final section of this appendix.

C.1. Fat-water phantom construction studies

Various studies have used a house-built fat-water phantom as verification for their research. Each uses their own materials and method. An overview of multiple studies and characteristics is given in Table C.1. Bush et al. present a detailed low cost, flexible and scalable protocol that gives guidance into creating a fat-water phantom [37]. Some of the other methods mentioned in Table C.1 are much alike.

Research	**No. vials**	**Volume per vial**	**Relevant concentrations**	**Fat substance**	**Emulsifying agent**	**Thickening agent**
Bush, 2018 [37]	5	100 mL	0% - 100%	Peanut oil	Tween and Span	Agar
Bernard, 2007 [38]	11	100 mL	0% - 100 %	Soybean oil	Sodium dodecyl sulphate	Carrageenan
Fischer, 2012 [39]	7	10 mL	0% - 50%	Goose fat	*unknown*	*unknown*
Bray, 2017 [40]	12	50 mL	0% - 70%	Peanut oil	Sodium dodecyl sulphate	Agar
Hines, 2009 [41]	11	40 mL	0% - 50%	Peanut oil	Sodium dodecyl sulphate	Agar
Bydder, 2007 [42]	7	55 mL	0% - 20%	Soybean oil	*unknown*	*unknown*

Table C.1: Overview of house-built phantoms and their characteristics.

C.2. Method

C.2.1. Construction

The house-built phantom was constructed based on the method of Bush et al. [37]. Their method allows for the use of simply basic lab supplies, such as a hotplate, beakers, etc, and easily accessible materials (distilled

water, agar, peanut oil, sodium benzoate and surfactants). The resulting phantom consists of homogeneous oil-in-water emulsions that remain stable during MR imaging and the FF value range they build is relevant for this research project. Some comments and adjustments are mentioned here:

- Peanut oil is used, because its proton NMR spectrum is similar to that of triglyceride protons in human adipose tissue [43].
- We expect to measure a fat fraction of 5-15% in relevant structures during a later experiment, hence a fat fraction of 15% is added to the the five other fractions described by Bush et al. (0%, 25%, 50%, 75% and 100%) [37].
- The volume of the vials will be reduced from 100 mL to 50 mL. The aim is to create a compact overall shape of the phantom, while the diameter of each vial is large enough to allow for informative measurements. The entire phantom has to fit under a head coil and be compact in order to minimally affect the magnetic field.

C.2.2. Image acquisition and analysis

The true FF values of the fat-water phantom vials are not precisely known, which means that the accuracy of the protocols described in FIXME cannot be evaluated carefully using this phantom. Only an estimation of the accuracy can be done by comparing the outcome values to the corresponding expected FF values of the phantom vials. The precision, however, can be checked using test-retesting. Multiple scans are performed to find out whether these yield the same results.

MR measurements are done using the protocols from Table 3.1 with a clinical 3T Philips Ingenia MRI system using a head-and-neck coil. The DICOM files are loaded into a Matlab script in order to analyse the data. A circular segmentation was made manually for each vial and converted into a mask. This mask was multiplied with multiple slices of each scan in which the segmentations would lay correctly in place of the vials. Then, the mean FF value and standard deviation for each vial is calculated based on all the pixels in the segmentation.

C.3. Results

C.3.1. Construction

The goal was to fill six vials with emulsions of 0%, 15%, 25%, 50%, 75% and 100% fat to water ratio. However, only five fat-water solutions were successfully built (Figure C.1). The fat-water combinations with a fat fraction higher than 30% failed, because the substances separated. This could be due to the high concentration of lipids against polar molecules (water). Furthermore, the temperature of both solutions was hard to regulate on only one hotplate. The water solution had probably cooled down a bit a few times, but the agar had never fully coagulated.

The desired FF values could differ a few percentage points from the true values, since part of the liquids lingered in the graduated cylinders while adding solutions to the Erlenmeyer flask. The water solution lingered more, so it is expected that the FF value of the vials is higher than attempted.

C.3.2. FF measurements

The first measurements on the house-built phantom show a relatively low standard deviation for the measured FF values in the chosen ROIs (see Table C.2). This could be an indication for both nice homogeneity of the phantoms and high precision in the measurements. The outcomes are quite consistent in repeated measurements and the surrounding does not seem to have a large influence on the measured FF values (Scan 3 was surrounded by air instead of water). There is quite a large difference between the outcomes of the protocols; the standard protocol gives higher FF values. The measurements on the Calimetrix phantom give more information on the accuracy of both protocols.

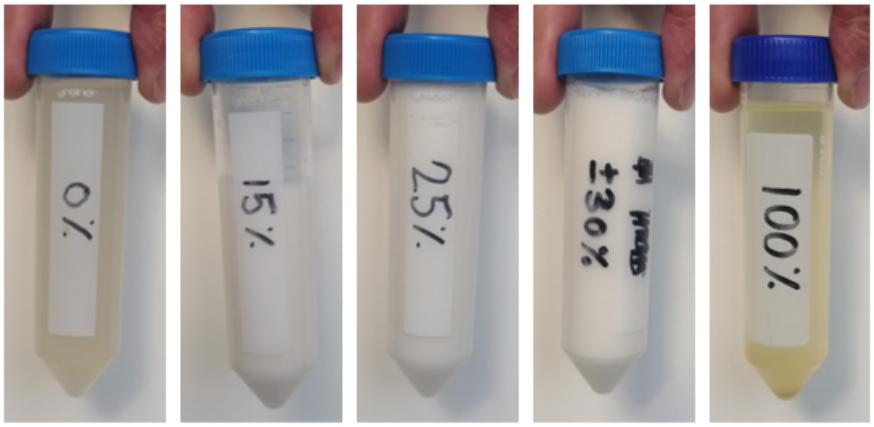


Figure C.1: Successfully built phantom vials with FF percentages mentioned on vial.

	Vial 1 0%		**Vial 2 15%**		**Vial 3 25%**		**Vial 4 30%**		**Vial 5 100%**	
Scan	mean	std	mean	std	mean	std	mean	std	mean	std
LUMC Muscle 1	1.5	1.0	11.2	2.1	17.7	1.0	22.2	1.3	101.0	1.2
LUMC Muscle 2	0.3	0.9	11.1	2.1	17.6	1.0	22.4	1.3	101.0	1.2
LUMC Muscle 3	0.7	1.2	10.1	2.0	20.0	0.8	21.3	0.9	101.8	0.5
LUMC Muscle 4	0.9	1.1	11.9	1.0	21.2	1.0	25.1	1.4	101.4	0.7
Average	**0.9**	**0.5**	**11.1**	**0.7**	**19.1**	**1.8**	**22.7**	**1.6**	**101.3**	**0.4**
Standard 5	-1.4	0.9	20.1	0.9	32.5	1.0	38.0	1.1	98.6	1.7
Standard 6	0.1	1.1	28.7	1.2	42.3	1.5	48.1	1.5	98.1	1.3
Average	**-0.7**	**1.0**	**24.4**	**6.0**	**37.4**	**6.9**	**43.0**	**7.1**	**98.4**	**0.3**

Table C.2: Mean FF values (%) and standard deviations of six scans done in two MRI sessions. During all scans except for scan 'LUMC Muscle 3', the phantom vials were surrounded by water (in a container). During scan 3, the vials were surrounded by air.

Bibliography

[1] Athanassios Argiris, Michalis V Karamouzis, David Raben, and Robert L Ferris. Head and neck cancer. *The Lancet*, 371(9625):1695–1709, 2008. ISSN 0140-6736. doi: https://doi.org/10.1016/S0140-6736(08)60728-X.

[2] Eleni M. Rettig and Gypsyamber D'Souza. Epidemiology of head and neck cancer. *Surgical Oncology Clinics of North America*, 24(3):379–396, 2015. ISSN 1055-3207. doi: https://doi.org/10.1016/j.soc.2015.03.001. Head and Neck Cancer.

[3] Netherland Comprehensive Cancer Center (IKNL). Figures on cancer. Utrecht. Accessed on August 3rd, 2021. URL `Availableathttp://www.cijfersoverkanker.nl/selecties/Dataset_1/img5528b4d09d2a2`.

[4] Sue P Humphrey and Russell T Williamson. A review of saliva: normal composition, flow, and function. *The Journal of prosthetic dentistry*, 85(2):162–169, 2001.

[5] Matthijs H Valstar, Bernadette S de Bakker, Roel JHM Steenbakkers, Kees H de Jong, Laura A Smit, Thomas JW Klein Nulent, Robert JJ van Es, Ingrid Hofland, Bart de Keizer, Bas Jasperse, et al. The tubarial salivary glands: A potential new organ at risk for radiotherapy. *Radiotherapy and Oncology*, 154:292–298, 2021.

[6] Antonetta C Houweling, Marielle EP Philippens, Tim Dijkema, Judith M Roesink, Chris HJ Terhaard, Cornelis Schilstra, Randall K Ten Haken, Avraham Eisbruch, and Cornelis PJ Raaijmakers. A comparison of dose–response models for the parotid gland in a large group of head-and-neck cancer patients. *International Journal of Radiation Oncology* Biology* Physics*, 76(4):1259–1265, 2010.

[7] Avraham Eisbruch, Marco Schwartz, Coen Rasch, Karen Vineberg, Eugene Damen, Corina J Van As, Robin Marsh, Frank A Pameijer, and Alfons JM Balm. Dysphagia and aspiration after chemoradiotherapy for head-and-neck cancer: which anatomic structures are affected and can they be spared by IMRT? *International Journal of Radiation Oncology* Biology* Physics*, 60(5):1425–1439, 2004.

[8] XiaoShen Wang and Avraham Eisbruch. IMRT for head and neck cancer: reducing xerostomia and dysphagia. *Journal of radiation research*, 57(S1):i69–i75, 2016.

[9] Terence T Sio, Huei-Kai Lin, Qiuling Shi, G Brandon Gunn, Charles S Cleeland, J Jack Lee, Mike Hernandez, Pierre Blanchard, Nikhil G Thaker, Jack Phan, et al. Intensity modulated proton therapy versus intensity modulated photon radiation therapy for oropharyngeal cancer: first comparative results of patient-reported outcomes. *International Journal of Radiation Oncology* Biology* Physics*, 95(4):1107–1114, 2016.

[10] Matteo Paoletti, Anna Pichiecchio, Stefano Cotti Piccinelli, Giorgio Tasca, Angela L Berardinelli, Alessandro Padovani, and Massimiliano Filosto. Advances in quantitative imaging of genetic and acquired myopathies: clinical applications and perspectives. *Frontiers in neurology*, 10:78, 2019.

[11] Timothy JP Bray, Manil D Chouhan, Shonit Punwani, Alan Bainbridge, and Margaret A Hall-Craggs. Fat fraction mapping using magnetic resonance imaging: insight into pathophysiology. *The British journal of radiology*, 91(1089):20170344, 2017.

[12] O Grundmann, GC Mitchell, and KH Limesand. Sensitivity of salivary glands to radiation: from animal models to therapies. *Journal of dental research*, 88(10):894–903, 2009.

[13] Keiko Teshima, Ryuji Murakami, Ryoji Yoshida, Hideki Nakayama, Akimitsu Hiraki, Toshinori Hirai, Yuji Nakaguchi, Naoko Tsujita, Etsushi Tomitaka, Mitsuhiro Furusawa, et al. Histopathological changes in parotid and submandibular glands of patients treated with preoperative chemoradiation therapy for oral cancer. *Journal of radiation research*, pages 1204060258–1204060258, 2012.

[14] Hing-Chiu Chang, Chun-Jung Juan, Hui-Chu Chiu, Yi-Jui Liu, Cheng-Chieh Cheng, Su-Chin Chiu, Cheng-Yu Chen, Guo-Shu Huang, and Hsiao-Wen Chung. Parotid fat contents in healthy subjects evaluated with iterative decomposition with echo asymmetry and least squares fat-water separation. *Radiology*, 267(3):918–923, 2013.

[15] Yoshitaka Kise, Toru Chikui, Yasuo Yamashita, Koji Kobayashi, and Kazunori Yoshiura. Clinical usefulness of the mDIXON Quant the method for estimation of the salivary gland fat fraction: comparison with MR spectroscopy. *The British journal of radiology*, 90(1077):20160704, 2017.

[16] Toru Chikui, Yasuo Yamashita, Yoshitaka Kise, Tomonori Saito, Kazutoshi Okamura, and Kazunori Yoshiura. Estimation of proton density fat fraction of the salivary gland. *The British journal of radiology*, 91(1085):20170671, 2018.

[17] Nan Zhou, Chen Chu, Xin Dou, Weibo Chen, Jian He, Jing Yan, Zhengyang Zhou, and Xiaofeng Yang. Early evaluation of radiation-induced parotid damage in patients with nasopharyngeal carcinoma by T2 mapping and mdixon quant imaging: initial findings. *Radiation Oncology*, 13(1):1–6, 2018.

[18] Guo-Yi Su, Chuan-Bing Wang, Hao Hu, Jun Liu, Hong-Yuan Ding, Xiao-Quan Xu, and Fei-Yun Wu. Effect of laterality, gender, age and body mass index on the fat fraction of salivary glands in healthy volunteers: assessed using iterative decomposition of water and fat with echo asymmetry and least-squares estimation method. *Dentomaxillofacial Radiology*, 48(3):20180263, 2019.

[19] Chen Chu, Qianqian Feng, Huayong Zhang, Shengnan Zhao, Weibo Chen, Jian He, Lingyun Sun, and Zhengyang Zhou. Evaluation of salivary gland fat fraction values in patients with primary sjögren's syndrome by mDIXON Quant imaging: Initial findings. *European journal of radiology*, 123:108776, 2020.

[20] Ianessa A Humbert, Scott B Reeder, Eva J Porcaro, Stephanie A Kays, Jean H Brittain, and JoAnne Robbins. Simultaneous estimation of tongue volume and fat fraction using IDEAL-FSE. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 28(2):504–508, 2008.

[21] Vinit Baliyan, Chandan J Das, Raju Sharma, and Arun Kumar Gupta. Diffusion weighted imaging: technique and applications. *World journal of radiology*, 8(9):785, 2016.

[22] DM Koh and AR Padhani. Diffusion-weighted MRI: a new functional clinical technique for tumour imaging. *The British Journal of Radiology*, 79(944):633–635, 2006.

[23] Venla Loimu, Tiina Seppälä, Mika Kapanen, Laura Tuomikoski, Heidi Nurmi, Antti Mäkitie, Mikko Tenhunen, and Kauko Saarilahti. Diffusion-weighted magnetic resonance imaging for evaluation of salivary gland function in head and neck cancer patients treated with intensity-modulated radiotherapy. *Radiotherapy and Oncology*, 122(2):178–184, 2017.

[24] Marion Smits. MRI biomarkers in neuro-oncology. *Nature Reviews Neurology*, pages 1–15, 2021.

[25] W Thomas Dixon. Simple proton spectroscopic imaging. *Radiology*, 153(1):189–194, 1984.

[26] Jingfei Ma. Dixon techniques for water and fat imaging. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 28(3):543–558, 2008.

[27] Scott B Reeder, Angel R Pineda, Zhifei Wen, Ann Shimakawa, Huanzhou Yu, Jean H Brittain, Garry E Gold, Christopher H Beaulieu, and Norbert J Pelc. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): application with fast spin-echo imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 54(3): 636–644, 2005.

[28] TG Perkins, A Duijndam, H Eggers, E de Weerdt, and YHE Rijckaert. Fat-free imaging. 2015.

[29] Houchun H Hu, Takeshi Yokoo, Mustafa R Bashir, Claude B Sirlin, Diego Hernando, Dariya Malyarenko, Thomas L Chenevert, Mark A Smith, Suraj D Serai, Michael S Middleton, et al. Linearity and bias of proton density fat fraction as a quantitative imaging biomarker: A multicenter, multiplatform, multivendor phantom study. *Radiology*, page 202912, 2021.

[30] Mikael Skorpil, Henric Rydén, Johan Wejde, Elisabet Lidbrink, Otte Brosjö, and Johan Berglund. The effect of radiotherapy on fat content and fatty acids in myxoid liposarcomas quantified by MRI. *Magnetic resonance imaging*, 43:37–41, 2017.

[31] Tim Schakel, Johannes M Hoogduin, Chris HJ Terhaard, and Marielle EP Philippens. Diffusion-weighted mri with minimal distortion in head-and-neck radiotherapy using a turbo spin echo acquisition method. *Medical physics*, 44(8):4188–4193, 2017.

[32] Fritz Schick. SPLICE: sub-second diffusion-sensitive mr imaging using a modified fast spin-echo acquisition mode. *Magnetic resonance in medicine*, 38(4):638–644, 1997.

[33] Harriet C Thoeny, Frederik De Keyzer, Chris Boesch, and Robert Hermans. Diffusion-weighted imaging of the parotid gland: Influence of the choice of b-values on the apparent diffusion coefficient value. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 20(5):786–790, 2004.

[34] Model 300 calimetrix PDFF phantom, https://www.calimetrix.com/pdff. URL `https://www.calimetrix.com/pdff`.

[35] Ruvini Navaratna, Ruiyang Zhao, Timothy J Colgan, Houchun Harry Hu, Mark Bydder, Takeshi Yokoo, Mustafa R Bashir, Michael S Middleton, Suraj D Serai, Dariya Malyarenko, et al. Temperature-corrected proton density fat fraction estimation using chemical shift-encoded MRI in phantoms. *Magnetic Resonance in Medicine*, 2021.

[36] Pilar Jimenez-Royo, Michele Bombardieri, Coziana Ciurtin, Michalis Kostapanos, Anwar R Tappuni, Natasha Jordan, Azeem Saleem, Teresa Fuller, Kathleen Port, Elena Pontarini, et al. Advanced imaging for quantification of abnormalities in the salivary glands of patients with primary sjögren's syndrome. *Rheumatology*, 60(5):2396–2408, 2021.

[37] Emily C Bush, Aliya Gifford, Crystal L Coolbaugh, Theodore F Towse, Bruce M Damon, and E Brian Welch. Fat-water phantoms for magnetic resonance imaging validation: a flexible and scalable protocol. *JoVE (Journal of Visualized Experiments)*, (139):e57704, 2018.

[38] Clare P Bernard, Gary P Liney, David J Manton, Lindsay W Turnbull, and Christian M Langton. Comparison of fat quantification methods: a phantom study at 3.0 T. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 27(1):192–197, 2008.

[39] Michael A Fischer, Dimitri A Raptis, Matteo Montani, Rolf Graf, Pierre-Alain Clavien, Daniel Nanz, Hatem Alkadhi, and Hans Scheffel. Liver fat quantification by dual-echo MR imaging outperforms traditional histopathological analysis. *Academic radiology*, 19(10):1208–1214, 2012.

[40] Timothy JP Bray, Alan Bainbridge, Shonit Punwani, Yiannis Ioannou, and Margaret A Hall-Craggs. Simultaneous quantification of bone edema/adiposity and structure in inflamed bone using chemical shift-encoded mri in spondyloarthritis. *Magnetic resonance in medicine*, 79(2):1031–1042, 2018.

[41] Catherine DG Hines, Huanzhou Yu, Ann Shimakawa, Charles A McKenzie, Jean H Brittain, and Scott B Reeder. T1 independent, T2* corrected mri with accurate spectral modeling for quantification of fat: validation in a fat-water-SPIO phantom. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 30(5):1215–1222, 2009.

[42] Mark Bydder, Takeshi Yokoo, Gavin Hamilton, Michael S Middleton, Alyssa D Chavez, Jeffrey B Schwimmer, Joel E Lavine, and Claude B Sirlin. Relaxation effects in the quantification of fat using gradient echo imaging. *Magnetic resonance imaging*, 26(3):347–359, 2008.

[43] Huanzhou Yu, Ann Shimakawa, Charles A McKenzie, Ethan Brodsky, Jean H Brittain, and Scott B Reeder. Multiecho water-fat separation and simultaneous R estimation with multifrequency fat spectrum modeling. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 60(5):1122–1134, 2008.