MRI-based methods for response assessment to neoadjuvant therapy in esophageal cancer

Methodes om met MRI de reactie van het oesofaguscarcinoom op neoadjuvante therapie te bepalen

Sophie Elise Heethuis

MRI-based methods for response assessment to neoadjuvant therapy in esophageal cancer

PhD Thesis, Utrecht University, The Netherlands

 Financial support for publication of this thesis was kindly provided by Elekta.

Cover design:	Robbie Rens
Lay-out:	Sophie Heethuis
Typeset:	LaTeX
Printed by:	Gildeprint
ISBN:	978-94-9301-449-7

MRI-based methods for response assessment to neoadjuvant therapy in esophageal cancer

Methodes om met MRI de reactie van het oesofaguscarcinoom op neoadjuvante therapie te bepalen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Utrecht
op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op donderdag 11 oktober 2018 des ochtends te 10.30 uur

door

Sophie Elise Heethuis
geboren op 20 mei 1990
te Alkmaar

Promotor: Prof. dr. ir. J.J.W. Lagendijk
Copromotor: Dr. ir. A.L.H.M.W. van Lier

Contents

Chapter 1

General introduction

Incidence of esophageal cancer

In the Netherlands over 2500 people are newly diagnosed with esophageal cancer every year (2016) [1]. Worldwide this number reaches 456.000 newly diagnosed cases each year (2012) with the highest rates in Eastern Asia and in Eastern and Southern Africa. It is the 8^{th} most common cancer worldwide and the incidence is rapidly increasing [2].

Esophageal cancer occurs most often in men (75%) and can be differentiated in two main tumor types; adenocarcinomas (AC) and squamous cell carcinomas (SCC). Specific risk factors for SCC are known to include heavy use of tobacco and alcohol, low socioeconomic status and poor oral hygiene. The occurrence of AC is highly associated with obesity and with gastroesophageal reflux disease (GERD) causing a Barrett's esophagus, in which metaplastic changes occur to the lower part of the esophagus due to chronic acid exposure [3–5]. In eastern countries (e.g. Central Asia, China) it is known that the majority (90%) of esophageal cancer cases are SCC. Whereas, in Western countries the incidence of SCC has steadily been decreasing over the last decades (reduction in alcohol/tobacco use) while the incidence of AC has increased rapidly (increased obesity prevalence and Barrett's esophagus) [2]. In the Netherlands only around 25% of the cases are SCC (2008) [6].

Dysphagia is one of the most common symptoms associated with esophageal cancer. Since difficulty in swallowing does not occur until the circumference of the esophagus is already narrowed one-third to one-half of the normal size, esophageal cancer is known to manifest itself late [7]. As a result, the disease is associated with poor 5-year overall survival rates of 15-25% [4]. In the Netherlands 1813 deaths associated to esophageal cancer occur yearly [1].

Staging and treatment of esophageal cancer

To confirm the suspicion of esophageal cancer an upper-gastrointestinal endoscopy combined with biopsy is used. Staging of esophageal tumors is performed according to the TNM classification system, based on the local invasion of the tumor (T-stage), lymph node involvement (N-stage) and spread to other organs (M-stage). The T-stage is divided in different stages, dependent on the depth of tumor growth (Table 1.1 and Figure 1.1) .

Most effective techniques used for TNM classification, determining loco-regional lymph node status and to assess loco-regional spread and distant metastases are both endoscopic ultrasonography (EUS) and ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}F-FDG-PET)/CT [8–10]. Depending on the TNM classification a curative or palliative treatment course is prescribed.

Table 1.1: T-stages for esophageal cancer.

T-stage	Depth of tumor growth
Tis	high-grade dysplasia
T1	tumor invades lamina propria (T1a) or submucosa (T1b)
T2	tumor invades muscularis propria
T3	tumor invades the outer layer of the esophagus (adventitia)
T4a	resectable tumor invading pleura, pericardium, or diaphragm
T4b	unresectable tumor invading other adjacent structures, such as the aorta, vertebral body, and trachea

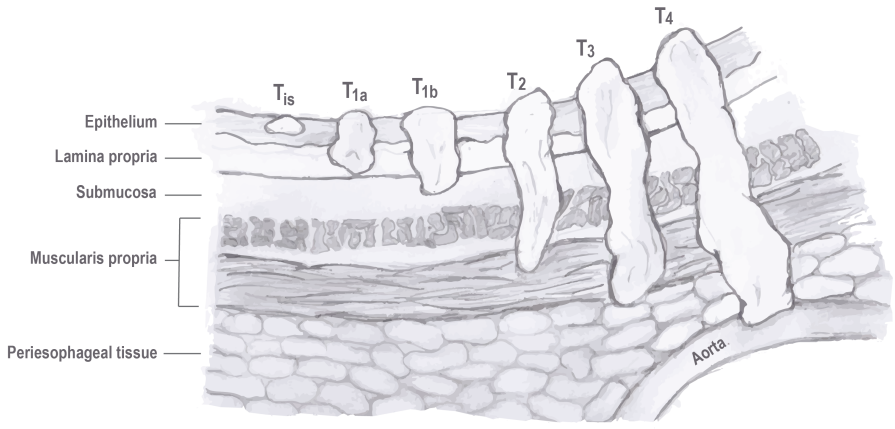


Figure 1.1: A schematic representation of the different T-stages for esophageal cancer, visualizing the differences in depth of tumor growth into the different layers.

Standard of care in the Netherlands for patients with advanced resectable non-metastatic esophageal cancer (i.e. T1N1, T2, T3 and T4a), in a curative setting, involves neoadjuvant chemoradiotherapy (nCRT) followed by surgery. nCRT consists of a radiotherapy schedule with the delivery of 41.4 Gy in 23 daily fractions of 1.8 Gy [11] and chemotherapy includes weekly intravenous administration of carboplatin and paclitaxel. Introduction of esophagectomy with en-bloc lymphadenectomy results in 5-year survival rates of 34% to 36% [11, 12]. Additional delivery of nCRT increases 5-year survival rates by approximately 13% as was consistently shown in recent trials and a meta-analysis [11,13,14]. Furthermore, no effect was seen by including nCRT prior to surgery on postoperative short-term health-related quality of life (HRQOL) domains compared with surgery alone [15]. Esophagectomy with en-bloc lymphadenectomy is known to be a technically challenging and major surgery with a risk of (postoperative) complications such as anastomotic leakage or pneumonia [11,16]. After surgery, pathologic assessment of the resection specimen is performed to determine the tumor regression grade (TRG). Classification is often performed according to the Mandard scoring system in which Mandard 1 represents no residual cancer and Mandard 5

absence of regressive changes [17].

Treatment of early-stage esophageal cancer (i.e. Tis and T1ab) with esophagectomy results in overall 5-year survival as high as 90% [18]. However, due to the disadvantages of esophagectomy (perioperative morbidity and mortality) early-stage esophageal cancer is increasingly treated using endoscopic submucosal dissection (ESD) or endoscopic mucosal resection (EMR) [19, 20]. Both techniques preserve the esophagus without compromising survival in a selected group of patients [21, 22].

Assessment of treatment response

Selecting the appropriate patient group for certain treatment strategies would enable individualized treatment. In case accurate prediction of good response to nCRT for patients with esophageal cancer would be available, a wait-and-see approach with omission of surgery and close clinical follow-up could be applied for the 29% of patients who reach a pathologic complete response (pCR) after nCRT [11]. On the contrary, poorly responding patients may benefit from treatment modification/intensification or primary surgery without nCRT. Furthermore, survival rates are known to be dependent on tumor-type, in which SCC has a large benefit of additional nCRT prior to surgery, whereas AC show less gain in survival [14]. Unfortunately, diagnostic modalities that are currently used to identify pathologic response yield unsatisfactory results. Highest pooled sensitivities and specificities were found for ^{18}F-FDG-PET(/CT), however, with values ranging from 67% to 70% this is still insufficient to justify changes in clinical decision-making [23, 24]. Since magnetic resonance imaging (MRI) can reach high soft-tissue contrast in contrary to other imaging modalities, MRI is increasingly used for visualization and delineation of tumors [25–28].

MRI for response monitoring

As described by Hanahan et al. cancer is believed to gain ten hallmark capabilities in order to emerge [29]. These capabilities like e.g. evading apoptosis, insensitivity for anti-growth signals and sustained angiogenesis induce abnormalities like increased microstructural density and the presence of e.g. distorted and enlarged vessels, erratic blood flow and leakiness of vessels. When neoadjuvant treatment is effective for example tumor cellularity is expected to decrease again. Since (functional) MRI serves the potential to measure deviations in micro-structure and functioning of the tissue it is increasingly researched in relation to detect changes during treatment and potentially predict response.

Publications concerning diffusion-weighted MRI (mapping microstructural density) and dynamic contrast-enhanced MRI (mapping tissue perfusion) and its possible applications to cancer already appeared over a decade ago [30, 31]. However, for

esophageal cancer research using quantitative MRI data is limited so far, since visualiziation of this tumor site was considered for a long time as technically challenging. Main challenges included organ motion, respiratory and cardiac action, and proximate considerable blood flow, leading to motion and flow artefacts [10]. Furthermore, the central location of the esophagus in the body leads to a relatively high noise level. The introduction of several technical innovations reduced the mentioned image artefacts initiating a new era of esophageal cancer research. Several (pilot) studies were published researching the feasibility of acquiring quantitative MRI data and the potential value in relation to therapy management, monitoring and prediction in patients with esophageal cancer [32–34].

Dynamic contrast-enhanced imaging
Tumors are characterized by angiogenesis paired with increased vascular permeability. Dynamic contrast-enhanced (DCE-) MRI imaging visualizes tissue perfusion, enabling the differentiation of tumorous tissue [35, 36]. In DCE-MRI fast T_1-weighted imaging is performed during and after the intravenous injection of a contrast agent (CA). Images are acquired sequentially every 2-3 seconds for typically a couple of minutes. After passing the lungs and the heart the CA reaches the tumor tissue. Due to the vascular permeability the CA leaks out of the capillaries into the extravascular extracellular space, which changes the T_1 of the tissue and elevating the signal intensity on the MR image. This direct relationship (for a spoiled gradient echo MR-sequence) between the MR-signal and the concentration of contrast agent is described in the following formulas:

$$S \propto sin(\alpha) \cdot \frac{1 - e^{-TR/T_1}}{1 - cos(\alpha) \cdot e^{-TR/T_1}} \tag{1.1}$$

$$\frac{1}{\Delta T_1} = r_1 \cdot [C] \tag{1.2}$$

in which α is the flipangle, TR the repetition time, r_1 and [C] the relaxivity constant and concentration of the contrast agent, respectively. $(1/\Delta T_1)$ reflects the change in relaxation rate after inflow of the contrast agent with respect to the true tissues relaxation rate.

In Figure 1.2c the calculated concentration of CA after injection is visualized as measured in the aorta, healthy tissue and tumorous tissue. The concentration as measured in the aorta shows a clear peak after injection, followed by a smaller peak after the second pass through the artery. Comparing the concentration inflow as measured in healthy tissue (liver) to tumorous tissue, the increased blood flow in the tumor is visible through a higher and steeper rise in signal followed by a higher tail of the curve due to the vascular permeability.

To perform quantitative analysis and enable comparison of scans originating from

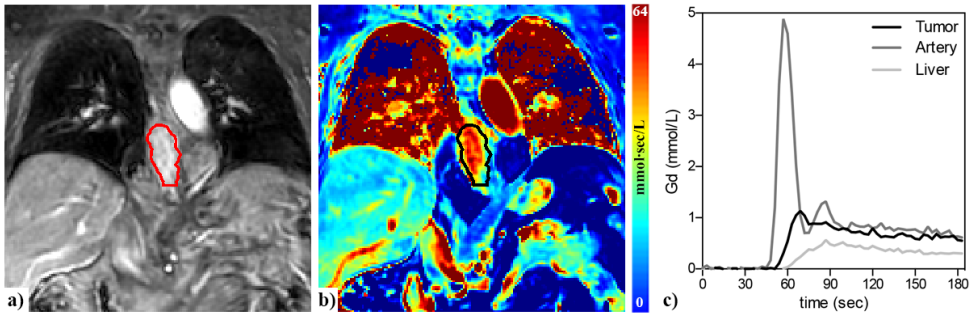


Figure 1.2: In **(a)** a coronal slice of the DCE-series (30th scan) is shown and in **(b)** the corresponding AUC_{60}-map (integration of 60 seconds after inflow of CA). In both images the tumor delineation is indicated (red and black, respectively). **(c)** Visualization of the differences in behavior of the concentration uptake after inflow of contrast between the artery, tumor and liver (i.e. healthy tissue).

different time points, scanners, different CA, a T_1-map prior to contrast injection can be used to calculate the concentration of CA. Performing an integration over the area-under-the-concentration (AUC) versus time curve, a higher accumulation of contrast agent is visible in the obtained AUC-map comparing the tumor to healthy tissue. In Figure 1.2b this is visible in an AUC-map in which an integration was performed of 60 seconds, after inflow of the contrast agent.

Diffusion-weighted imaging
Diffusion-weighted (DW-) MRI visualizes the variations in free diffusion (i.e. Brownian motion) of water molecules between different tissues. By measuring at different diffusion strengths (b-values), the apparent diffusion coefficient (ADC) can be deduced to quantify the differences in microstructural density of tissues (Equation 1.3).

$$S_b = S_0 \cdot e^{-b \cdot ADC} \tag{1.3}$$

in which S_b represents the signal at a specific b-value, S_0 the signal intensity at $b = 0\,s\,mm^{-2}$ and ADC the apparent diffusion coefficient.

In tissue with high cellularity (i.e. tumorous tissue) the signal decay is hampered resulting in high signal intensities in images with high b-values [37, 38]. Since the ADC is inversely correlated with tissue cellularity, tumors appear hyperintense on scans with high b-values (Figure 1.3c), but hypointense on ADC maps (Figure 1.3d).

It is known that the signal decay is not purely mono-exponential but that it is influenced by perfusion in the capillary network, which elevates estimated ADC values if not taken into account (Figure 1.4). Intravoxel incoherent motion (IVIM) is a method in which not only molecular diffusion of water is taken into account, but also the microcirculation of blood in the capillary network. To be able to separate the

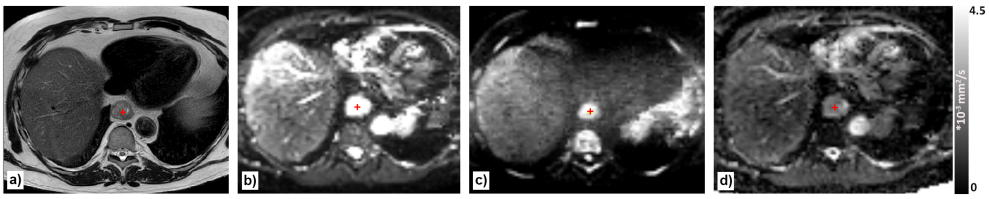


Figure 1.3: In **(a)** a transverse anatomical T2-weighted MRI scan is shown. In **(b)** and **(c)** images with b-values of 0 and 800 s mm^{-2} are shown, respectively. The corresponding ADC-map is shown in **(d)**. All four images are linked to each other and the tumor is indicated by a red crosshair.

contribution of perfusion, a lot of scans with varying b-values are acquired, especially multiple scans with a b-value < 200 s mm^{-2}. A well-known model used for analysis is a bi-exponential model introduced in 1988 by Le Bihan et al. [39]. This is a two-compartment model assuming the presence of a diffusion-dominated compartment and a perfusion-dominated compartment. As model parameters a perfusion-dominated pseudo-diffusion (D^*) and a diffusion-dominated slow diffusion (D) are determined, together with a volume of blood fraction (f) (Equation 1.4, Figure 1.4).

$$S_b = S_0 \cdot [(1-f) \cdot e^{-b \cdot D} + f \cdot e^{-b \cdot D^*}] \tag{1.4}$$

For fitting this model often first a mono-exponent is fitted to b-values ≥ 200 s mm^{-2}, to determine D and f, followed by fitting Equation 1.4 with D and f as input parameters. However, multiple other approaches to fit this model have been researched [40, 41], one of which is for example the Bayesian-Probability approach. In this approach a prior probability is assumed for the different model parameters. After each iteration, providing new knowledge, the posterior distribution becomes the next prior distribution. Finally, computing the location of the maxima in the probability density functions of D, D^* and f estimates the three model parameters.

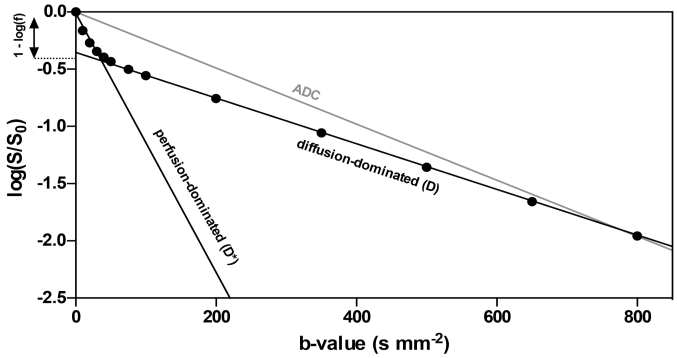


Figure 1.4: Typical decay of a signal from Equation 1.4, in which the perfusion-dominated part is visible in the low b-values. The grey line indicates a mono-exponential ADC following Equation 1.3 (with b-values of 0 and 800 s mm^{-2}).

Multimodality assessment of response

Since different modalities visualize different tumor characteristics they have the potential to provide complementary information when predicting response. Using a model in which multiple modalities or parameters are used to predict response has the potential to increase the predictive values which are currently reached. Therefore, the use of multiparamatric models combining two or more modalities is increasingly researched for a variety of tumor sites [42–45].

For a personalized treatment approach several strategies could be used to alter the 'standard treatment' as explained previously, depending on patient outcome. In a strategy in which one wants to apply an organ-preserving wait-and-see approach (omit surgery) you want your prediction model to have a high positive predictive value (PPV), since you want to be sure to only include complete (or near-complete) responders. On the other hand, in a strategy in which the neoadjuvant treatment is modified or even terminated to initiate surgery sooner, you only want to include non-responders and therefore a prediction model with a high negative predictive value (NPV) is required. Figure 1.5 represents a possible personalized treatment strategy in which, dependent on predicted response, several treatment paths can be chosen.

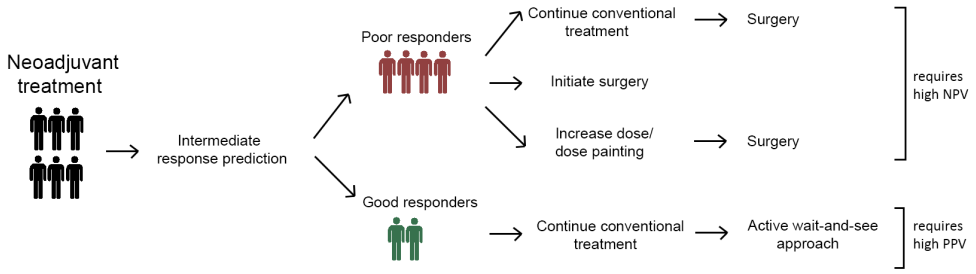


Figure 1.5: A flowchart of a possible personalized treatment strategy in which depending on response prediction different treatment paths can be followed.

Tumor motion

Several studies have shown that the motion in esophageal tumors can be quite extensive and depend on e.g. tumor stage and tumor location [46–49]. In radiotherapy, in current clinical practice, margins are applied to account for setup errors and uncertainties in the position of the target volume due to organ motion [50]. Because of these margins, for most patients adequate dose coverage will be obtained in the target. However, using margins also introduces larger toxicity to surrounding healthy tissue (see Figure 1.6) which in turn increases the risk of adverse effects induced by

radiotherapy [51]. The desired ratio between the probability of tumor cell kill versus the probability of normal tissue complications is also dependent on whether the treatment is curative or palliative.

Knowledge of the tumor motion during radiotherapy could enable more conformal dose delivery and individualized treatment strategies. For this purpose systems were developed using X-ray imaging (CyberKnife) [52] or internal marker tracking (Calypso) [53] during delivery. Moreover, an integrated MRI and Linac [54–57] provides the advantages of an MR system, allowing precise real-time tracking of both tumor and surrounding soft tissues during dose delivery. Ultimately, these techniques have the potential to reduce currently used motion margins, decreasing irradiation to surrounding healthy tissues and allow optimization of the dose to the tumor.

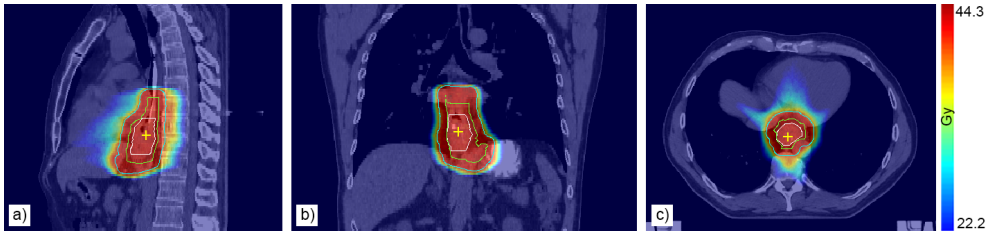


Figure 1.6: An example of a radiotherapy planning using volumetric modulated arc radiotherapy (VMAT). In **(a)** a sagittal slice of a CT scan is shown with the gross tumor volume (GTV) delineated in white (39.5 mL). The clinical target volume (CTV) includes an extra 3 cm cranio-caudal direction and 0.5 cm circumferential (adjusted for anatomical structures) and is shown in green (114.7 mL). An additional 1 cm margin for the planning target volume (PTV) is indicated in light-blue (353.7 mL). The planned dose for delivery is shown as overlay in color, in which only values $\geq$50% of the maximum dose of 44.3 Grey are shown. **(b)** and **(c)** represent the coronal and transverse view, respectively, in which it is clearly visible that proximate organs at risk (lungs and heart) are also receiving radiation. The yellow crosshair indicates identical position in all three planes.

Outline thesis

This thesis mainly focuses on the use of functional MR imaging to predict response to nCRT for patients with esophageal cancer. Using the pathological assessment of the resection specimen as ground truth of response to nCRT, enables us to quantify the strength of the different modalities to predict response.

Chapter 2 covers the use of quantitative DCE-MRI for response prediction to nCRT. The potential of DW-MRI as quantitative method for response prediction in esophageal cancer has already been demonstrated. Since ^{18}F-FDG-PET/CT is often used in the clinic, its correlation with DW-MRI is dealt with in **Chapter 3**. The complementary value of both DCE-MRI and DW-MRI in response prediction to nCRT for patients with esophageal cancer is discussed in **Chapter 4**. Additionally, in view of our future large multi-center trial for response assessment in patients with esophageal cancer, including IVIM-MRI, in **Chapter 5** a methodical study is presented researching the influence of using three fitting approaches for estimating IVIM-MRI parameters in patients with esophageal cancer.
As mentioned previously, the introduction of the MRLinac in the clinic brings high potential enabling real-time tumor tracking during irradiation and potentially reduce current used motion margins. As a first step for esophageal cancer to benefit from this new system, the variation in motion within a fraction and throughout the course of treatment, as determined on cine-MRI, is discussed in **Chapter 6**.

Finally, a general discussion of the main results of this thesis followed by future perspectives are presented in **Chapter 7**.

Chapter 2

Dynamic contrast-enhanced MRI for treatment response assessment in patients with esophageal cancer receiving neoadjuvant chemoradiotherapy

The following chapter is based on:

Sophie E. Heethuis, Peter S.N. van Rossum, Irene M. Lips, Lucas Goense, Francine E. Voncken, Onne Reerink, Richard van Hillegersberg, Jelle P.Ruurda, Marielle E. Philippens, Marco van Vulpen, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier

Radiotherapy and Oncology (2016) Vol. 120 (1), p.128-135

Abstract

Purpose
To explore and evaluate the potential value of dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) for the prediction of pathologic response to neoadjuvant chemoradiotherapy (nCRT) in esophageal cancer.

Material and methods
Twenty-six patients underwent DCE-MRI before, during (week 23) and after nCRT, but before surgery (pre/per/post, respectively). Histopathologic tumor regression grade (TRG) was assessed after esophagectomy. Tumor area-under-the-concentration time curve (AUC), time-to-peak (TTP) and slope were calculated. The ability of these DCE-parameters to distinguish good responders (GR, TRG 1-2) from poor responders (noGR, TRG$\geq$3), and pathologic complete responders (pCR) from no-pCR was assessed.

Results
Twelve patients (48%) showed GR of which 8 patients (32%) pCR. Analysis of AUC change throughout treatment, $AUC_{per\text{-}pre}$, was most predictive for GR, at a threshold of 22.7% resulting in a sensitivity of 92%, specificity of 77%, PPV of 79%, and a NPV of 91%. $AUC_{post\text{-}pre}$ was most predictive for pCR, at a threshold of -24.6% resulting in a sensitivity of 83%, specificity of 88%, PPV of 71%, and a NPV of 93%. TTP and slope were not associated with pathologic response.

Conclusions
This study demonstrates that changes in AUC throughout treatment are promising for prediction of histopathologic response to nCRT for esophageal cancer.

Introduction

Worldwide, esophageal cancer is the eight most common cancer and the incidence rate is rapidly increasing [2]. Esophageal cancer has a poor prognosis with 5-year overall survival rates ranging from 15 to 25% [4]. Esophagectomy with en-bloc lymphadenectomy for patients with resectable non-metastatic disease results in 5-year survival rates of 34% to 36% [11, 12]. Neoadjuvant chemoradiotherapy (nCRT) increases these rates by approximately 13% as was consistently shown in recent trials and a meta-analysis [11, 13, 14]. Therefore nCRT is currently considered as the standard treatment with curative intent for both adenocarcinomas (AC) and squamous cell carcinomas (SCC). However, not all patients benefit equally, as patient outcome depends heavily on the response to chemo(radio)therapy [13, 58, 59]. In 29% of the patients pathologic complete response (pCR) to nCRT is found, with increased 5-year overall survival rates up to 48-65% [11, 58, 60, 61].

With accurate response prediction before surgery the treatment strategy could potentially be improved. Depending on patient outcome, adaptive approaches could be explored such as an organ-preserving wait-and-see approach, modification of nCRT or termination of neoadjuvant therapy to initiate surgery sooner [62].

Endoscopic biopsy and/or ultrasonography and ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}F-FDG-PET) have been extensively studied for neoadjuvant treatment response assessment in esophageal cancer [23, 24, 63–65]. Unfortunately, these imaging modalities yield insufficient accuracies in the prediction of pathologic response [27, 66]. Among the studied modalities ^{18}F-FDG-PET seems the best so far. Two meta-analyses on treatment response monitoring, found that scans from the first two weeks of nCRT were most predictive for pathologic response, showing similar or superior diagnostic accuracy as opposed to pre- and post-treatment scanning only [23, 24, 67]. Accordingly, investigators proposed that the second week might be optimal, because significant tumor regression can already be found in responders while the image interpretation may not yet be influenced by radiation esophagitis [24]. Overall pooled sensitivities and specificities were, however, still insufficient to justify changes in clinical decision-making with values ranging from 67% to 70% [23, 24]. Also non-image based methods using molecular biomarkers and clinical parameters have shown potential for response prediction [68–70].

Recently, diffusion-weighted magnetic resonance imaging (DW-MRI) was described as a potential method for treatment response monitoring and prediction in esophageal cancer [34, 66]. It was reported that the change in apparent diffusion coefficient (ADC) measured prior and during nCRT is predictive for response. The reported sensitivity and specificity were higher than for the aforementioned methods, showing the potential of MRI-based techniques for early treatment response assessment and prediction.

Pilot studies on dynamic contrast-enhanced (DCE-)MRI demonstrated the feasibility for esophageal cancer imaging [32, 71]. However, the use of DCE-MRI to measure (early) response of esophageal tumors to nCRT with histopathology as reference standard has not previously been described. Therefore, the purpose of this study was to investigate the potential of nonparametric analysis of DCE-MRI for the (early) prediction of pathologic response to nCRT in patients with esophageal cancer. Since DCE-MRI visualizes different physiological properties (perfusion and vascular permeability in tumor microenvironment) compared to DW-MRI (diffusion in tissues), it could potentially provide complementary information beyond DW-MRI. In this exploratory study not only pre- and post-treatment measurements were obtained, but also approximately two weeks after initiation of treatment, enabling comparison of the predictive potential of perfusion parameters at three different time points.

Materials and Methods

Study population

This prospective study was approved by the local medical ethical committee and written informed consent was obtained from all patients. Patients with contraindications for MRI with contrast agent were not eligible for inclusion. Six patients who participated in the study but did not receive all three MRI scans due to patients wish of discontinuation (n=5) or urgent non-elective surgery (n=1), were excluded. In addition, one patient was excluded due to a technical failure during MR acquisition. Twenty-six consecutive patients with biopsy-proven esophageal cancer who received nCRT followed by surgery, and completed all three MRI studies in our institutes (University Medical Center Utrecht [UMCU], n=19; and the Netherlands Cancer Institute [NKI], n=7) from August 2013 to January 2015, were included. One patient with severe pneumonia on the initial MRI, with an infiltrate adjacent to the tumor, was excluded for further analysis.

All patients received five weeks of neoadjuvant treatment, involving weekly intravenous administration of carboplatin (area under the curve of 2 mg/mL per minute) and paclitaxel $50\,\mathrm{mg\,m^{-2}}$ (50 $\mathrm{mg/m^2}$ body-surface area) with concurrent radiotherapy (41.4 Gy in 23 fractions of 1.8 Gy) [11]. Five to ten weeks after completion of nCRT all patients underwent esophagectomy with en-bloc two-field lymphadenectomy and gastric conduit reconstruction with cervical anastomosis. After resection, pathologic assessment of the resection specimen included determination of the tumor regression grade (TRG) according to Mandard, using an identical protocol in both institutes [17].

Table 2.1: Patient and treatment-related characteristics.

	Study population UMCU	Study population NKI/AVL
Gender		
Male	17	6
Female	1	1
Age, years (at start RT)	65.1 ± 7.5	59.5 ± 13.4
Clinical T-stage		
T2	4 (22%)	1 (14%)
T2-3	3 (17%)	0 (0%)
T3	11 (61%)	6 (68%)
Clinical N-stage		
N0	6 (33%)	4 (57%)
N1	3 (17%)	2 (29%)
N2	8 (44%)	1 (14%)
N3	1 (5.5%)	0 (0%)
Type		
SCC	1 (5.5%)	2 (29%)
AC	16 (89%)	5 (71%)
ASC	1 (5.5%)	0 (0%)
Location		
Middle third of oesophagus	1 (5.5%)	1 (14%)
Distal third of oesophagus	13 (72%)	3 (43%)
Gastroesophageal junction	4 (22%)	3 (43%)
Acquisition, nr of days		
Before (before start nCRT)	5.7 ± 2.9	4.4 ± 3.6
During (after start nCRT)	11.3 ± 2.6	12.9 ± 3.0
After (after completion nCRT)	38.7 ± 13.7	46.3 ± 10.2

SCC = squamous cell carcinoma ; AC = adenocarcinoma ; ASC = adenosquamous carcinoma.

MRI acquisition

MR images were acquired at the following three time points: prior to treatment (pre), after 8-13 fractions nCRT (per) and 3-9 weeks after completion of treatment, prior to surgery (post). All MR images were acquired with 1.5 T systems Philips Achieva or Philips Ingenia (Best, the Netherlands), using the Torso coil (16 channel) or Anterior/Posterior (28 channel) receive coils, respectively. For anatomical verification, a T_2-weighted scan was performed with a multi-slice turbo spin echo sequence (TR/TE = 1983/100 ms, resolution = 0.67x0.67x4 mm^3), using a navigator for respiratory triggering [66]. A DCE-MRI serie of 62 images was acquired using a three-dimensional spoiled gradient echo sequence (TR/TE = 3.43/1.53 ms, flip angle = 20°, matrix size = 432x432x33, reconstructed image voxel size = 1.18x1.18x3 mm^3), with a 3-second

interval. After the 10^{th} image, the contrast agent (CA) gadobutrol (Gd-BT-DO3A, Gadovist; Schering AG, Berlin, Germany) was injected at a dose of 0.1 mmol/kg of body weight with an automatic syringe pump at a flow rate of 1 ml/sec followed by saline injection. The scanned volume included the heart in order to prevent artefacts in the aorta due to pulsatile flow. In both institutes the same imaging protocol was used. However, at the NKI a different CA was used, Dotarem (Gadoteric acid, 0.5mM; Guerbet, Paris, France), with a fixed dose of 7.5 mmol for each patient.

Prior to the dynamic series, five acquisitions for varying flip angles ($\alpha = 2°/6°/10°/12°/16°$) with identical scanning properties were acquired for determination of pre-contrast T_1 values. This flip angle series was chosen to be sensitive to a large range of tissue T_1 values. Additionally, DW-MRI scans were acquired with b-values of 0, 200 and 800 $s\,mm^{-2}$ (STIR fat suppression, resolution = 3.5x3.5x4 mm^3) [66].

Image processing

Delineation of the primary tumor on the T_2-weighted scans was divided over two clinicians (P.S.N.v.R and I.M.L.). The delineation was adapted in the scans obtained during and after nCRT to account for tumor alterations. For definition of the cranio-caudal tumor length the DW-MRI with b = $800\,s\,mm^{-2}$ was used. A radiation oncologist (O.R.), with over 10 years of experience, verified all delineations. To account for breathing motion within the DCE-MRI series, scans were rigidly registered to a scan after contrast-enhancement, which led to the best retrospective motion compensation [72] (Supplementary Figure 2.1). Finally, the delineation was cropped with an isotropic margin of 2 mm to account for the conversion from T_2-weighted scan to the DCE-MRI, residual motion and partial volume effects.

Image analysis

In order to enable comparison of scans between different sessions in time, independent of MRI scaling settings, institute, MRI scanner or CA, the image intensity was converted to concentration. For this purpose, T_1 pre-contrast relaxation times were calculated with in-house developed software using multiple flip angle sequences. Relaxivity values of 4.7 and 3.6 L $mmol^{-1}$ s^{-1} for Gadovist and Dotarem, respectively, were used [73].

A nonparametric approach was chosen for the analysis of DCE-MRI throughout time, as studies in different tumor sites indicate that this approach has an increased prognostic ability compared to parametric measures [74]. No analysis was performed per tumor subtype, as the subgroups of SCC and adenosquamous carcinoma (ASC) were too small. For analysis, the initial area-under-the-concentration versus time curve (AUC) was calculated (Supplementary Figure 2.1). The AUC reflects blood flow, vascular permeability and the fraction of interstitial space [75]. In addition, it is a relatively straight-forward model-free parameter, which can provide information about

tumor vascular changes. The time-to-peak (TTP) and slope of the concentration-time curve were mapped for each tumor, using a customized sigmoid-curve fitting method in Matlab (The Mathworks Inc, Natick, MA). To account for variations in the moment of contrast injection, the TTP was calculated with respect to the initial arterial CA increase, measured as the TTP in the artery (TTP_{aorta}) (Supplementary Figure 2.1).

The AUC was defined as the trapezoidal integral over the concentration-time curve over a period of 60 seconds after inflow of CA, based on TTP_{aorta}. Within the tumor delineation three percentiles, 25^{th} (P25), median (P50) and 75^{th} (P75), of the AUC were calculated. Additionally, relative differences between time points were calculated with respect to the first MRI scan session (pre):

$$AUC_{t\text{-}pre} = \frac{AUC_t - AUC_{pre}}{AUC_{pre}} \cdot 100\% \tag{2.1}$$

with t the second or third scan session (per or post).

Statistical analysis

This pilot study was of exploratory nature, therefore no formal sample size calculation has been performed. All statistical tests were performed using SPSS Statistics version 22 (IBM Corp., Armonk, NY, USA). Two different approaches were chosen. First, a distinction was made between a group of pathologic good responders (GR), defined as pCR (TRG 1) or near-pCR (TRG 2), and a group of poor responders (noGR) with TRG 3 or higher [17]. Second, analysis was performed on pCR (TRG 1) versus no-pCR (TRG$\geq$2). For both approaches the Mann-Whitney U test was used to compare the various parameters derived from DCE-MRI between the groups (if not stated otherwise). A p-value of <0.05 was considered statistically significant. Receiver operating characteristic (ROC) curve analysis was performed. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were estimated for an optimal threshold in which sensitivity and specificity were given equal weight. Correlation between AUC values and TRG scores were assessed by Spearmans rank correlation tests.

Results

In the 25 patients included, assessment of the histopathologic tumor type revealed in 21 patients (84%) AC, in three patients (12%) SCC and in one patient (4%) ASC of the oesophagus. Patient and tumor-related characteristics for both institutes are presented in Table 2.1. The pathology specimen showed a pathologic good response (i.e. TRG 1-2) to nCRT in 12 of 25 patients (48%), whereas the remaining 13 patients

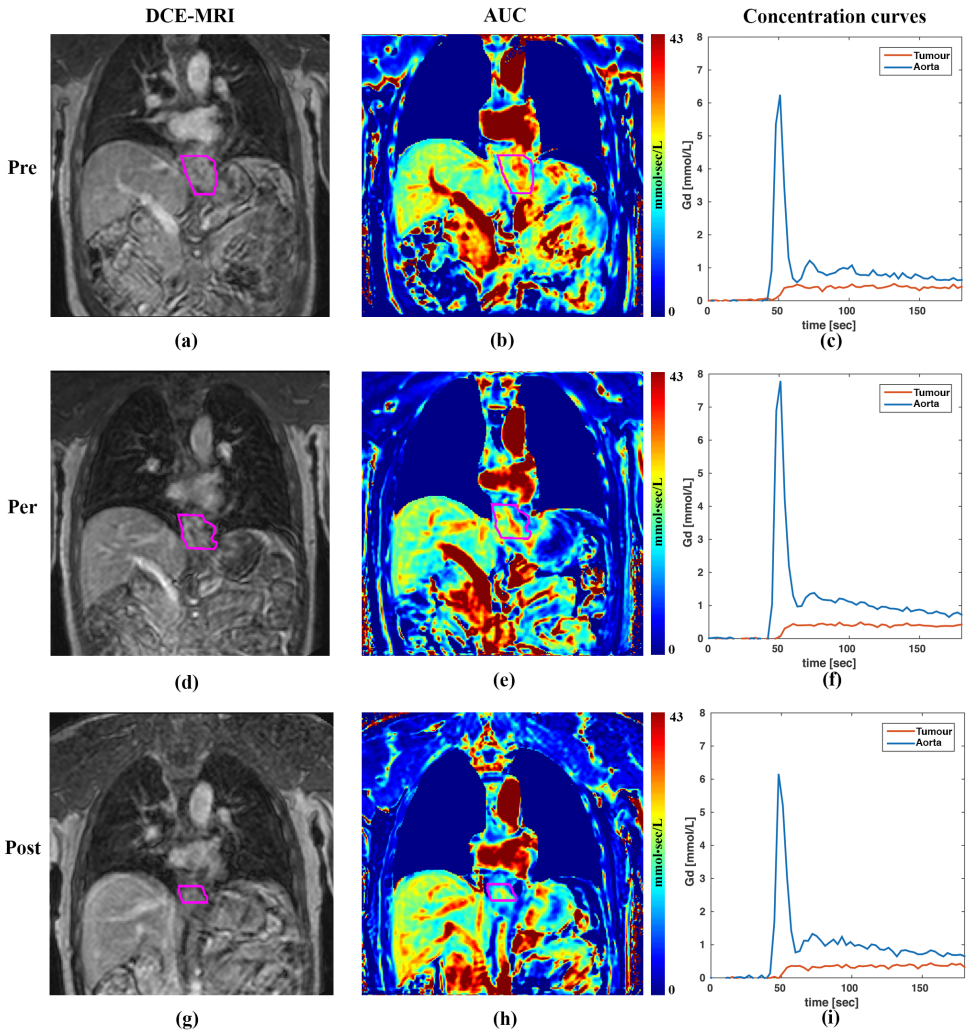


Figure 2.1: In **(a)** a coronal view of a DCE-MRI (after contrast agent injection) of a pathologic complete responder is shown before treatment (pre). The tumor, as delineated by a clinician based on the corresponding T_2-weighted scan, is shown in pink. In **(b)** the corresponding AUC map (60 seconds after contrast inflow) is shown. In **(d)** and **(e)** the same data is shown after two weeks of nCRT (i.e. during treatment). Comparing figs. **(b)** and **(e)**, a reduction is visible in AUC values during treatment. As shown in **(g)** and **(h)**, a further decrease in AUC values is visible after treatment as along with a decrease in volume. In all AUC maps a lung mask was used. Figs. **(c)**,**(f)** and **(i)** show the concentration time curve measured in a slice in the aorta and tumor, for pre-, per- and post-treatment scans, respectively. Clearly visible is that the inflow of contrast in the tumor closely follows the inflow in the aorta.

(52%) had a poor pathologic response (i.e. TRG ≥3). Among the good responders, 8 patients (32%) showed a pCR (i.e. TRG 1), and 4 patients (16%) had a near-pCR (i.e. TRG 2).

The tumor showed a rapid initial contrast-enhancement followed by a plateau over the first minutes (Figure 2.1). The timing of the enhancement closely followed the first-pass peak as observed in the aorta. This reflects the direct vascular connection between the aorta and the esophageal tumor. Small variations in peak height of the arterial input function (AIF) between different time points occurred regularly (also visible in Figure 2.1). However, in three patients with noGR, of which two showed an $AUC_{per\text{-}pre}$ >140%, an extreme increase was found in peak height in the second acquisition. In general, the AUC maps showed a heterogeneous uptake within the delineated tumor (Figure 2.1). In three patients, the delineation in the post-treatment scan was too small in cranio-caudal direction to subtract the margin and were therefore excluded from the analysis of this time point.

The association of the measured DCE-MRI parameters with GR versus noGR are presented in Table 2.2. Regarding the analysis of the AUC values for single time points, only the initial scan showed significant differences. The AUC for all patients at all time points is shown in Figure 2.2. The P25 AUC_{pre} was most significant, showing higher AUC in GR compared to noGR. The TTP and slope were not associated with response. The analysis of changes during treatment, showed in several noGRs an initial increase in AUC after 2 weeks of treatment followed by a decrease, while this was less apparent for GRs (Figure 2.2). Comparing the relative changes with respect to the initial scan session (Equation 2.1), for GR versus noGR, P75 $AUC_{per\text{-}pre}$ was found to be most significant showing higher AUC for noGR (Table 2.2/Figure 2.3a). The parameters determined at the final scan session before surgery were found to be

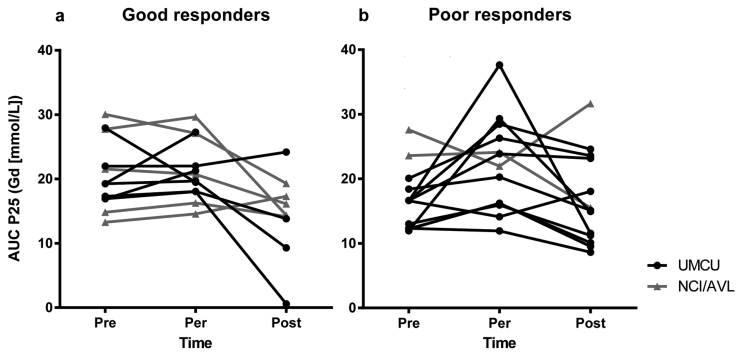


Figure 2.2: Course of P25 AUC values over time for pathologic good **(a)** and poor **(b)** responders. Every line represents a patient, with patients from the UMCU indicated by black circles and patients from the NKI/AVL by grey triangles. In case the delineation was too small, as explained in the results, the time point was excluded (3 post-treatment scans).

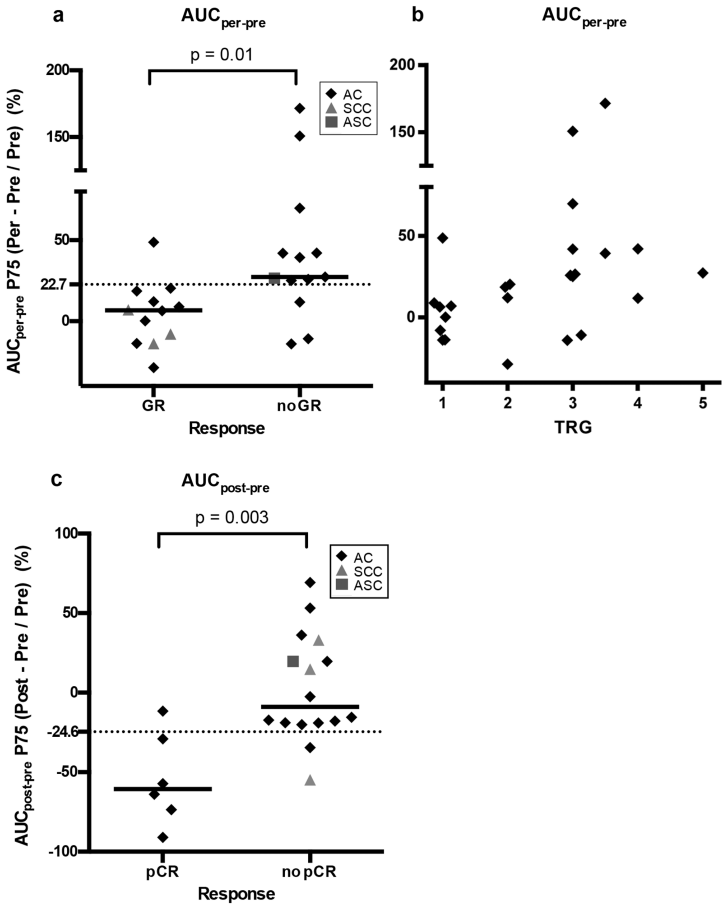


Figure 2.3: In **(a)** P75 $AUC_{per\text{-}pre}$ as function of response, good versus poor response. Each data point represents a patient and median values of both groups are indicated with solid lines. The dashed line represents the optimal threshold value of 22.7%. The legend specifies the tumor type for each patient (AC, SCC or ASC). In **(b)** P75 $AUC_{per\text{-}pre}$ is shown as a function of TRG score. Every data point represents a patient. In two patients the Mandard score was not conclusive (TRG 3 or 4), represented by a TRG of 3.5. **(c)** the P75 $AUC_{post\text{-}pre}$ is shown as function of response, pathologic complete response versus no-pathologic complete response. Median is indicated with a solid line for both groups and the threshold of -24.6% is indicated with a dashed line.

less associated with pathologic response. The optimal threshold for P75 $AUC_{per\text{-}pre}$ was determined at 22.7%, yielding a sensitivity of 92%, specificity of 77%, PPV of 79% and a NPV of 91%. Comparing P75 $AUC_{per\text{-}pre}$ as a function of the separate TRG scores, a moderate positive correlation was found (Spearmans rank order $\rho = 0.53$; $p = 0.006$) (Figure 2.3b).

Focussing on the previous significant parameters in the analysis of pCR versus no-pCR changes throughout treatment, several parameters also showed significant difference.

The associations of these parameters are presented in Table 2.3. The P75 $AUC_{post\text{-}pre}$ was found to be most predictive showing lower AUC for pCR compared to no-pCR (Figure 2.3c). The optimal threshold of -24.6% yielded a sensitivity of 83%, specificity of 88%, PPV of 71% and a NPV of 93%. The P75 $AUC_{per\text{-}pre}$ was also found to be significant, for an optimal threshold of 10.3%, a sensitivity of 88%, specificity of 82%, PPV of 70% and a NPV of 93% was reached. No significant differences were found in the TTP and slope, as well as in the analysis of single time point AUC.

Discussion

This prospective study demonstrates the potential of DCE-MRI for treatment response assessment and prediction to nCRT for patients with esophageal cancer. The change during treatment in P75 $AUC_{per\text{-}pre}$ was significantly different between GR and noGR with good diagnostic performance. This is particularly promising for future considerations regarding early modification or discontinuation of nCRT in anticipated noGRs. Additionally, the change in P75 $AUC_{post\text{-}pre}$ differed significantly between pCR and no-pCR, which could potentially aid in the clinical decision-making regarding the omission of surgery.

This is the first report on (early) response assessment measured by DCE-MRI with pathologic reference in esophageal cancer. In a previous study, a significant change in DCE-MRI parameters was observed in 12 patients with SCC of the oesophagus [32]. However, this feasibility study only performed a limited analysis (before vs. after treatment, single slice analysis, no comparison with pathologic response was given) compared to our study. The initial scan session showed predictive value for treatment outcome. However, acquiring a second scan after 2 weeks of nCRT added additional diagnostic value compared to pre- and post-treatment scanning only, resulting in higher predictive values. The finding that the change in tumor parameters early during nCRT is predictive for the pathologic outcome in terms of response, is in line with a previous study performed by our group using DW-MRI [66] as well as in other tumor sites using DCE-MRI [76, 77]. The large range in the timing of acquisition of the last scan, together with the fact that tumor delineations after chemoradiotherapy can be challenging, might have influenced the predictive values of the post-treatment scan to an unknown extent. The influence of this variability deserves attention in further studies.

For future studies on therapy adaptation, based on MRI response monitoring, it is advantageous that response can be measured early, on which adaptation could potentially be made in the first stages of neoadjuvant treatment. Furthermore, the scan prior to surgery showed diagnostic value in differentiation between patients with pCR and no-pCR, which could potentially be used in addition to the early differentiation between good and poor response. Equal weight was given to both sensitivity

Table 2.2: Association between (changes in) DCE-MRI parameters and pathologic good response (GR) versus poor response (noGR). For single time point AUC values, only highest significant parameters are presented. Data are presented as mean ± SD.

		Time	GR (n=12)	noGR (n=13)	p-value	AUC-ROC
AUC [mmol L^{-1} sec]	P25	Pre	20.6 ± 5.4	16.8 ± 4.8	**0.03**	0.76
		Per	21.2 ± 4.6	22.0 ± 7.3	1.00	0.50
		Post	14.3 ± 6.6*	16.7 ± 7.1	0.70	0.56
	ΔMedian(%)	Per-Pre	4.3 ± 16.3	41.0 ± 54.0	**0.02**	0.78
		Post-Pre	-30.8 ± 34.6*	2.0 ± 26.0	0.05	0.75
	ΔP75 (%)	Per-Pre	4.9 ± 20.1	46.7 ± 55.5	**0.01**	0.80
		Post-Pre	-37.1 ± 41.8*	4.0 ± 32.5	**0.03**	0.78
	ΔP25 (%)	Per-Pre	5.3 ± 17.5	36.8 ± 50.7	0.08	0.71
		Post-Pre	-26.0 ± 30.8*	-0.09 ± 20.8	0.05	0.75
Slope [mmol L^{-1} sec^{-1}]	Median	Pre	0.18 ± 0.07	0.13 ± 0.08	0.09	0.71
		Per	0.15 ± 0.06	0.17 ± 0.14	0.69	0.55
		Post	0.09 ± 0.05*	0.10 ± 0.09	0.90	0.52
	ΔMedian (%)	Per-Pre	-11.2 ± 38.8	34.4 ± 75.1	0.12	0.69
		Post-Pre	-56.3 ± 25.0*	3.7 ± 80.7	0.11	0.71
TTP [sec]	Median	Pre	7.8 ± 8.0	9.3 ± 5.9	0.43	0.60
		Per	9.3 ± 5.2	12.4 ± 9.9	0.69	0.55
		Post	15.9 ± 13.5*	17.6 ± 15.5	0.79	0.53
	ΔMedian (%)	Per-Pre	76.0 ± 160.5	30.6 ± 65.0	0.77	0.54
		Post-Pre	239.3 ± 329.9*	415.7 ± 1042.4	0.13	0.70

*n=9 instead of 12 for post scans due to the exclusion of 3 patients as explained in the results.

Table 2.3: Association between (changes in) DCE-MRI parameters and pathologic good response (GR) versus poor response (noGR). For single time point AUC values, only highest significant parameters are presented. Data are presented as mean ± SD.

		Time	pCR (n=8)	no-pCR (n=17)	p-value	AUC-ROC
AUC [mmol L^{-1} s]	ΔMedian (%)	Per-Pre	6.33 ± 15.4	31.41 ± 50.8	0.14	0.69
		Post-Pre	-42.57 ± 26.1*	0.28 ± 28.2**	**0.01**	0.84
	ΔP75 (%)	Per-Pre	4.47 ± 20.1	37.05 ± 52.3	**0.04**	0.76
		Post-Pre	-54.42 ±29.2*	2.82 ± 33.7**	**<0.01**	0.90
	ΔP25 (%)	Per-Pre	10.17 ± 16.39	27.10 ± 48.1	0.59	0.57
		Post-Pre	-36.25 ± 26.5*	-1.09 ± 22.3**	**0.01**	0.85

For the post-treatment scans due to the exclusion of 3 patients as explained in the results:
* n = 6 instead of 8
** n = 16 instead of 17.

and specificity, but depending on the purpose of the analysis (e.g. changing therapy, omitting surgery) this weight could be varied. In literature conflicting reports are found with respect to patient outcome and TRG 1 or 2. In terms of 5-year survival rates, some studies report statistically significant differences between patients with TRG 1 versus TRG≥2 [58, 60, 78], while in another study survival curves were found to be similar [79]. At this moment, it is debatable if the presented separation based on DCE-MRI between pCR/TRG≥2 or GR/noGR is sufficiently strong to support the decision to omit surgery. Therefore, the use of DCE-MRI (together with other imaging modalities) for treatment adaptations is subject of our on-going research. In addition to studying the complementary values of DW-MRI and DCE-MRI, it will become increasingly important to study the complementary values of clinical parameters, other imaging modalities (e.g. FDG-PET), and molecular biomarkers [68–70]. In this light it is promising that in rectal cancer omission of surgery or delayed resection in selected patient groups did not increase the risk of recurrent disease or affect survival compared to immediate surgery [80, 81].

The tumor contrast uptake curve pattern observed in this study resembles the curve described by Chang et al. [71], although the AIF differs. The observed AIF, showing a rapid in- and outflow of contrast, is more in line with the standard AIF as reported by Parker et al. [82]. To our knowledge, the trend showing lower $AUC_{\text{per-pre}}$ values for GR as opposed to noGR has not been reported previously for esophageal cancer. In a review by Li et al., summarizing various tumor sites, it was found that chemotherapy responders showed a drop in AUC, while radiotherapy responders showed an initial increase in AUC in the first 1-2 weeks, followed by an AUC drop [83].

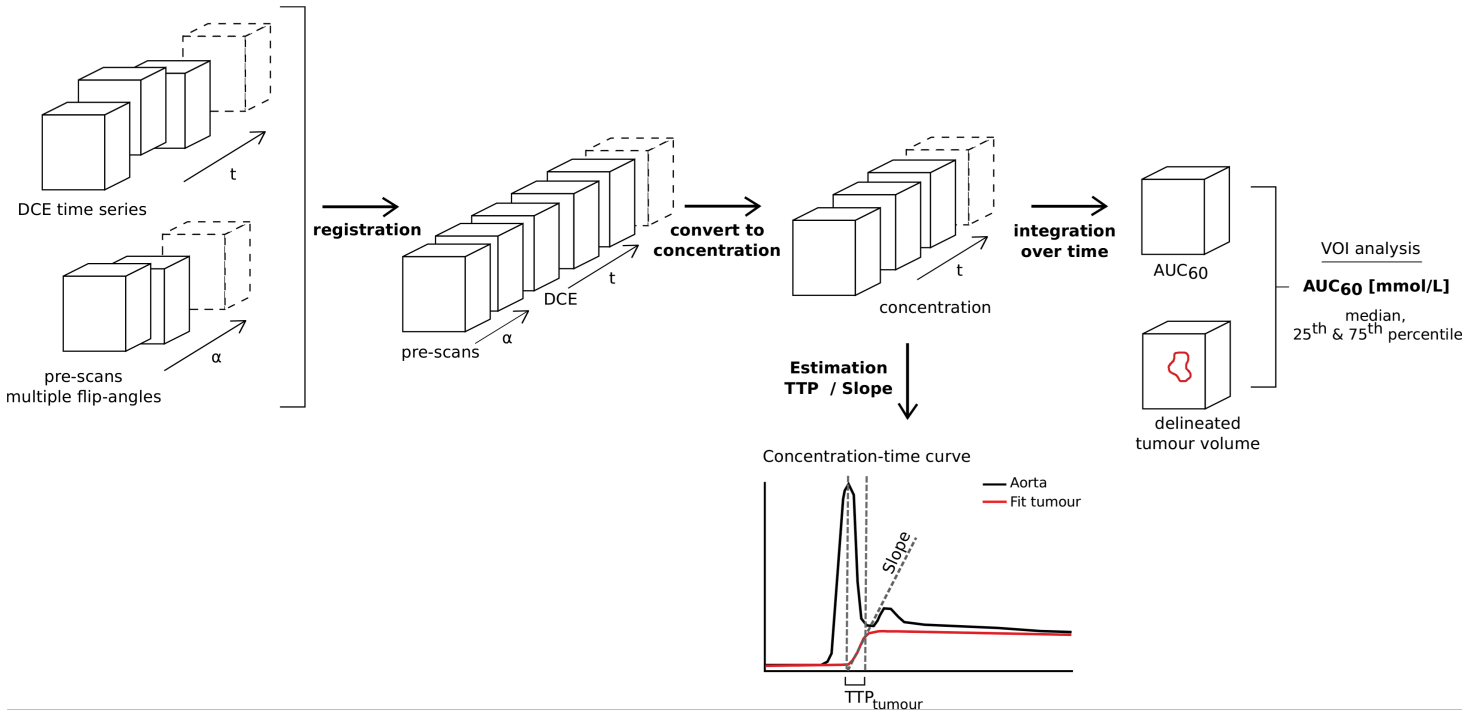
In this study a nonparametric approach was chosen, as it is not sensitive to errors arising from inaccurate determination of the AIF. Parametric approaches can be prone to errors due to uncertainty in input parameters, more complex fitting methods and lower reproducibility [74, 83–85]. On the other hand, Roberts et al. [84] suggested that the reproducibility of AUC and K_{trans} were similar, and therefore the parametric measure may be superior providing greater physiological insights. In future studies we intend to acquire multiple baseline acquisitions to gain more insight into the true variations and reproducibility of the AIF, to judge whether the use of a parametric model is justified.

As aforementioned, the AUC reflects blood flow and vascular permeability, which is increased in tumor tissue. The fact that P75 AUC resulted in superior predictive values could indicate that it is more sensitive to active (i.e. better perfused or more permeable) areas within the tumor. This could implicate that the delineated tumor is heterogeneous, and thus includes less active regions. In future research we intend to research the possibilities for voxel-based analysis of DCE-MRI.

Some limitations apply to this study. First, the small sample size of this exploratory

study may have caused both over- and underestimation of the true predictive values. Therefore, our conclusions require validation in a larger patient cohort. Second, a possible sensitivity of the AUC for the peak height of the AIF was found, which will be investigated by determining the reproducibility of DCE-MRI. This will also give further insight into the robustness of the method. Next, to compensate for organ motion due to respiration, we adopted a retrospective registration strategy. This improved geometrical overlap between consecutive scans. However, minor motion artefacts were visible in the individual scans. Strategies to compensate for motion prospectively are therefore expected to improve overall quality.

In conclusion, our study showed that analysis of DCE-MRI has potential for treatment response assessment and prediction in patients with esophageal cancer undergoing chemoradiotherapy. Changes in AUC within the tumor throughout the first two to three weeks of treatment were found to differ significantly between good and poor responders potentially allowing for early prediction of treatment response. In addition, the change in AUC prior to surgery was found to differ significantly between pathologic complete responders and non-pathologic complete responders.



Supplementary Figure 2.1: A flowchart of the registration and calculation of the concentration in the DCE-MRI time series. First, rigid registration is performed on the DCE-series as well as on the pre-treatment scans. Second, image intensities are converted to contrast agent concentration values. Conversion to concentration values enables comparison of different acquisitions. For analysis, the AUC over 60 seconds is calculated voxel wise, in which tumor median, 25^{th} and 75^{th} percentiles are estimated. TTP and slope are estimated using the concentration-time curves per voxel.

Chapter 3

Correlation between functional imaging markers derived from diffusion-weighted MRI and 18F-FDG-PET/CT in esophageal cancer

The following chapter is based on:

Lucas Goense, Sophie E. Heethuis, Peter S.N. van Rossum, Francine E. Voncken, Jan J.W. Lagendijk, Marnix G.E.H. Lam, Chris H. Terhaard, Richard van Hillegersberg, Jelle P.Ruurda, Stella Mook, Astrid L.H.M.W. van Lier, Steven H. Lin and Gert J. Meijer

Nuclear Medicine Communications (2018) 39(1):6067

Abstract

Objective
Both the apparent diffusion coefficient (ADC) acquired by diffusion-weighted magnetic resonance imaging (DW-MRI) and the standardized uptake value (SUV), acquired by ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG-PET/CT), are well-established functional parameters in cancer imaging. Currently it is unclear whether these two markers provide complementary prognostic and predictive information in esophageal cancer. The aim of this study was to evaluate the correlation between ADC and SUV in patients with esophageal cancer.

Methods
This prospective study included 76 patients with histologically proven esophageal cancer who underwent both DW-MRI and ^{18}F-FDG-PET/CT examinations before treatment. The minimum and mean ADC values (ADC_{min} and ADC_{mean}) of the primary tumor were assessed on MRI. Similarly, the glucose metabolism was evaluated by the maximum and mean SUV (SUV_{max} and SUV_{mean}) in the same lesions on ^{18}F-FDG-PET/CT images. Spearmans rank correlation coefficients were used to assess the correlation between tumor ADC and SUV values.

Results
The tumor ADC and SUV values as measures of cell density and glucose metabolism, respectively, showed negligible non-significant correlations (ADC_{min} versus SUV_{max}: r = -0.087, p = 0.457, ADC_{min} versus SUV_{mean}: r = -0.105, p = 0.369, ADC_{mean} versus SUV_{max}: r = -0.99, p = 0.349, ADC_{mean} versus SUV_{mean}: r = -0.111, p = 0.340). No differences in tumor ADC and SUV values were observed between the different histologic tumor types, stages and differentiation grades.

Conclusions
This study indicates that tumor cellularity derived from DW-MRI and tumor metabolism measured by ^{18}F-FDG-PET/CT are independent cellular phenomena in newly diagnosed esophageal cancer. Therefore, tumor ADC and SUV values may have complementary roles as imaging markers in the prediction of survival and evaluation of response to treatment in esophageal cancer.

Introduction

Esophageal cancer continues to affect more than 450.000 people yearly, and is the sixth leading cause of cancer-related mortality worldwide [4]. Currently, surgical resection of the esophagus combined with neoadjuvant therapy is the cornerstone of curative treatment for patients with non-metastasized esophageal cancer [4, 13, 14]. Definitive concurrent chemoradiotherapy is the preferred approach for inoperable locally advanced esophageal cancer [86, 87].

Esophageal cancer is usually diagnosed by endoscopy with biopsy and additional multimodality imaging is applied for staging. Since its clinical introduction, whole-body ^{18}F-fluorodeoxyglucose positron emission tomography with integrated computed tomography (^{18}F-FDG-PET/CT) has emerged as useful adjunct to conventional pre-treatment staging modalities for esophageal cancer [88]. The use of the ^{18}F-FDG tracer allows for quantitative assessment of increased cellular glucose metabolism in tumors by measuring the standardized uptake value (SUV). The SUV is the most commonly used quantitative index of ^{18}F-FDG uptake and is often expressed as its mean (SUV_{mean}) or maximum (SUV_{max}). In a pre-treatment setting the most important value of ^{18}F-FDG-PET/CT in the management of esophageal cancer lies in its ability to detect distant metastases and regional lymphadenopathy [9,89,90]. In addition, ^{18}F-FDG-PET/CT has shown potential for the prediction of long-term survival and evaluation of response to treatment [23, 24].

Meanwhile, diffusion-weighted magnetic resonance imaging (DW-MRI) is emerging as an advanced imaging technique in esophageal cancer imaging [27,33,91–93]. DW-MRI provides functional information based on the variation in (Brownian) motion of water molecules which is a marker for tissue density [38]. The apparent diffusion coefficient (ADC) is a quantitative measure for this variation and inversely correlates with tumor cellularity in various tumors [38], including esophageal cancer [94]. Furthermore, recent exploratory studies have shown that DW-MRI appears to provide valuable information regarding the assessment of response to treatment in esophageal cancer and in several other malignancies [66, 95–97].

Both ADC and SUV values are established parameters in cancer imaging, although they may refer to different aspects of tumor pathophysiology. Currently, it is unclear whether these two imaging markers provide complementary information with regard to the prediction of survival and evaluation of response to treatment in esophageal cancer. Therefore, the purpose of this prospective study was to assess the correlation between SUV and ADC values in order to evaluate whether these values are correlated or independent functional imaging markers in newly diagnosed esophageal cancer.

Materials and Methods

Data were gathered from a multicenter prospective study in which three international cancer centers participated, including The University Medical Center Utrecht (UMCU), The Netherlands Cancer Institute (NKI-AVL), and The University of Texas MD Anderson Cancer Center (MDACC). This prospective study was approved by the institutional review boards of each center separately and written informed consent was provided by all patients. The study was conducted in accordance with the Health Insurance Portability and Accountability Act (HIPAA) and was registered with ClinicalTrials.gov, number NCT02125448.

Study population

Patients presented at the three participating centers from October 2013 to November 2015 with newly diagnosed biopsy-proven esophageal cancer planned to receive neoadjuvant chemoradiotherapy followed by surgery were eligible for inclusion (Table 3.1). Exclusion criteria included a history of thoracic radiotherapy, contraindications for MRI, or more than 30 days interval between MRI and ^{18}F-FDG-PET/CT imaging. The diagnostic work-up consisted of an endoscopy with biopsy for diagnosis, endoscopic ultrasound, and integrated ^{18}F-FDG-PET/CT scan for clinical staging.

MR image acquisition

Patients underwent MRI scanning with anatomical T_2-weighted and DW-MRI sequences within two weeks prior to nCRT. The MRI examinations were either performed on a 1.5 Tesla scanner equipped with a 16 or 28-element phased-array receive coil for thoracic imaging (UMCU and NKI-AVL; Achieva or Ingenia; Philips Medical Systems, Best, The Netherlands) or on a 3.0 Tesla scanner equipped with a 32-channel torso phased array coil (MDACC; Discovery, GE Medical Systems, Milwaukee [WI], USA). The MRI scan protocol was specifically developed for esophageal cancer patients and made similar across the three participating centers prior to initiation of the current study [91]. Patients were scanned in supine position without administration of anti-peristaltic agents. Transverse DW-images were obtained under free breathing conditions with the following scan parameters: repetition time 7503 ms, echotime 180 ms, slice thickness 4 mm, 50 slices, field of view 260x560 mm, voxel size 3.25 mm, total acquisition duration was approximately 6 minutes. Three b-values were acquired: 0, 200 and 800 $\mathrm{s\,mm^{-2}}$. Sagittal and transverse T_2-weighted images were obtained with a navigator that monitors the position of the diaphragm using a fast intertwined 1D-MRI acquisition, in order to trigger scanning exclusively during the end of the expiration [98].

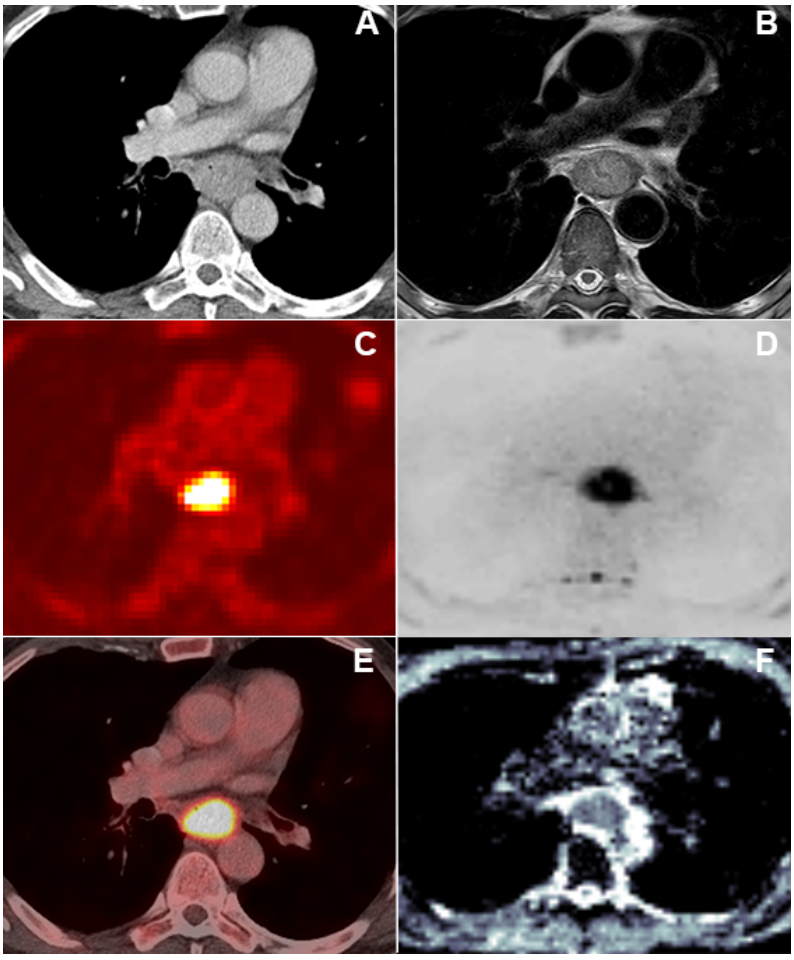


Figure 3.1: Example of a mid-esophageal adenocarcinoma of a 72 year old male patient. CT **(A)** demonstrates a round-shaped esophageal tumor with high uptake on ^{18}F-FDG-PET **(C)** and the fused images **(E)**. Corresponding tumor on T_2-weighted imaging **(B)** with high signal on DWI-MRI (b=800 s mm^{-2}) **(D)** and corresponding ADC map **(F)** with restricted diffusion at the location of the tumor.

^{18}F-FDG-PET/CT image acquisition

The ^{18}F-FDG-PET/CT examinations were performed on dedicated PET/CT systems (UMCU: mCT, Siemens, Erlangen, Germany; NKI-AVL: Gemini TF, Philips, Cleveland, Ohio; MDACC: Discovery RX, ST, or STE; GE Medical Systems, Milwaukee [WI], USA). Patients underwent injection of ^{18}F-FDG after fasting for at least six hours. Before injection of FDG, blood glucose levels were checked in every patient to excluded hyperglycemia. The activity of intravenously administered ^{18}F-FDG ranged between 190-370 MBq. Imaging started 60-90 minutes after administration of

^{18}F-FDG with a CT for attenuation correction with the following settings: 120 kV, 20-300 mAs, 0.5-s tube rotation time, pitch of 0.813-1.375 and 3-3.75 mm slice width. Following CT, PET scanning was performed from thigh to the base of skull in three-dimensional (3D) acquisition mode with 2-5 minutes per bed position. Images were reconstructed using iterative 3D-reconstruction.

Image analysis

All images were quantitatively assessed by authors (LG, SEH, PSNvR) who were trained in DW-MRI and ^{18}F-FDG-PET/CT image analysis. These reviewers were blinded to the clinical data, histopathologic results and the ADC maps when measuring ^{18}F-FDG-PET/CT images and vice versa. DW-MRI analysis including primary tumor delineation was performed using image analysis software package ITK-SNAP (version 3.4.0) [98]. The primary tumor was delineated on the DW-MR images with a b-value of $800\,\mathrm{s\,mm^{-2}}$ using automatic contouring, based on 2 clusters in pre-segmentation mode and default evolution parameters. The DW-MRI images were evaluated quantitatively by calculation of the ADC values for each voxel ($ln(S_b) = ln(S_0) - b \cdot ADC$, where S_0 is the signal intensity without diffusion-weighting) with b-values of 0, 200, and $800\,\mathrm{s\,mm^{-2}}$ [38, 66]. Subsequently, the volumes of interest (VOI) on the ADC maps were manually edited by two readers in consensus (LG and SEH) to ensure that only tumor tissue was covered. From the outlined VOIs, the mean and minimum ADC values per VOI were extracted. The mean ADC ($\mathrm{ADC_{mean}}$) was defined as the average ADC value of all voxels in each VOI, and minimum ADC ($\mathrm{ADC_{min}}$) as the lowest 2% percentile.

Based on the reconstructed ^{18}F-FDG-PET/CT images the primary tumor volumes were delineated using a semi-automatic gradient-based delineation method followed by manual editing by one reader in commercially available software (MIM Software, Cleveland, OH, USA). This method has been validated in a multi-observer study reporting superior accuracy, consistency and robustness compared with manual and threshold methods [99]. The ^{18}F-FDG uptake in the esophageal tumors as registered on the PET/CT images were evaluated using the SUV. SUV was defined as the ratio of tissue radioactivity concentration (Bq/mL) and the injected activity (Bq) divided by the body weight (kg$\times$1000). The software automatically calculated the metabolic tumor volume (in $\mathrm{cm^3}$), maximum and mean SUV ($\mathrm{SUV_{max}}$ and $\mathrm{SUV_{mean}}$) for each VOI.

Statistical analysis

To assess the strength and direction of correlation between DW-MRI metrics ($\mathrm{ADC_{min}}$, $\mathrm{ADC_{mean}}$) and ^{18}F-FDG-PET/CT ($\mathrm{SUV_{max}}$, $\mathrm{SUV_{mean}}$) values Pearson and Spearman correlation coefficients were calculated in case of parametric and non-parametric continuous variables, respectively. Pearson and Spearmans correlation coefficients

Table 3.1: Patient and tumor-related characteristics

Characteristic	n(%)
Gender	
Male	65 (86%)
Female	11 (14%)
Age, years[a]	61.4 ± 9.4
Histologic tumor type	
Adenocarcinoma	60 (79%)
Squamous cell carcinoma	13 (17%)
Other	3 (4%)
Histologic tumor grade	
Moderate differentiation	43 (57%)
Poor differentiation	33 (43%)
Tumor location	
Proximal third	2 (3%)
Middle third	9 (12%)
Distal third	49 (65%)
Gastro-esophageal junction	16 (20%)
Clinical T-stage	
cT2	12 (15%)
cT3	62 (82%)
cT4	2 (3%)
Clinical N-stage	
N0	25 (33%)
N1	30 (40%)
N2	20 (26%)
N3	1 (1%)
Interval PET/MRI in days[b]	17 (0-30)
ADC_{mean} ($10^{-3}mm^2\,s^{-1}$)[a]	2.05 ± 0.6
ADC_{min} ($10^{-3}mm^2\,s^{-1}$)[a]	0.75 ± 0.4
SUV_{mean}[a]	0.77 ± 0.4
SUV_{max}[a]	1.60 ± 1.0

[a] Data presented as mean ± standard deviation.
[b] Data presented as median with range.

may be interpreted as follows: a positive or negative correlation coefficient of 0.90-1.00 is considered very high; 0.70-0.89, high; 0.40-0.69, moderate; 0.30-0.49, low; and 0-0.29, negligible [100]. BlandAltman analysis was used for the assessment of agreement of volume parameters obtained from DW-MRI images and ^{18}F-FDG-PET/CT images [101]. Also, the tumor ADC and SUV were compared between squamous cell carcinomas and adenocarcinomas, between moderately and poorly differentiated tumor, and between different clinical T and N-stages using the Students T-test or

MannWhitney U test, for parametric and non-parametric variables, respectively. Statistical analysis was performed using SPSS 23.0 (IBM Corp. Armonk, NY, USA) and GraphPad Prism 6.07 software (GraphPad Software, La Jolla California USA). A p-value of <0.05 was considered statistically significant.

Results

Patient and tumor characteristics

In the study period, a total of 81 patients with newly diagnosed esophageal cancer who underwent both MRI and ^{18}F-FDG-PET/CT were potentially eligible for inclusion. Of these patients, 5 were excluded because the interval between MRI and ^{18}F-FDG-PET/CT imaging was more than 30 days (n=4) or the ADC map was of poor quality (n=1). The final study population comprised of 76 patients with a mean age of 61.4 years (standard deviation [SD], 9 years), and 65 (86%) of them were male. Histologic tumor types included adenocarcinoma (n=60, 79%), squamous cell carcinoma (n=13, 17%) or other types (n=3, 4%). The distribution of the esophageal tumor types corresponds with those of western populations. Patient and tumor-related characteristics are presented in Table 3.1.

Tumor ADC and SUV values

The mean ±SD of the ADC_{mean} and ADC_{min} obtained from the 76 esophageal tumors were 2.05 $\pm 0.6 \cdot 10^{-3}$ $mm^2\,s^{-1}$ and 0.75 $\pm 0.04 \cdot 10^{-3}$ $mm^2\,s^{-1}$, respectively. The average tumor volume as determined by semi-automatic delineation on DW-MRI was 32.6 ±21.4 cm^3 (range, 7.2-85 cm^3). The mean ±SD of the SUV_{mean} and SUV_{max} from ^{18}F-FDG-PET/CT were 7.7 ±0.4 and 16.0 ±1.0, respectively. Average metabolic tumor volume measured by semi-automatic ^{18}F-FDG-PET/CT contouring was 25.2 ±21.7 cm^3 (range, 2.9-94.0). Figure 3.1 demonstrates an example of ^{18}F-FDG-PET/CT and DW-MRI images in one of the patients.

Correlation analysis of ADC and SUV values

The tumor ADC and SUV values showed negligible non-significant correlations (ADC_{min} versus SUV_{max}: r=-0.087, p=0.457, ADC_{min} versus SUV_{mean}: r=-0.105, p=0.369, ADC_{mean} versus SUV_{max}: r=-0.99, p=0.349, ADC_{mean} vs. SUV_{mean}: r=-0.111, p=0.340). Figure 3.2 shows the scatter plots of the correlations between the different studied tumor ADC and SUV values.

Comparison between tumor volumes

In Bland-Altman analysis the mean of the difference between DW-MRI-based and ^{18}F-FDG-PET-based tumor volumes (and corresponding 95% limits of agreement)

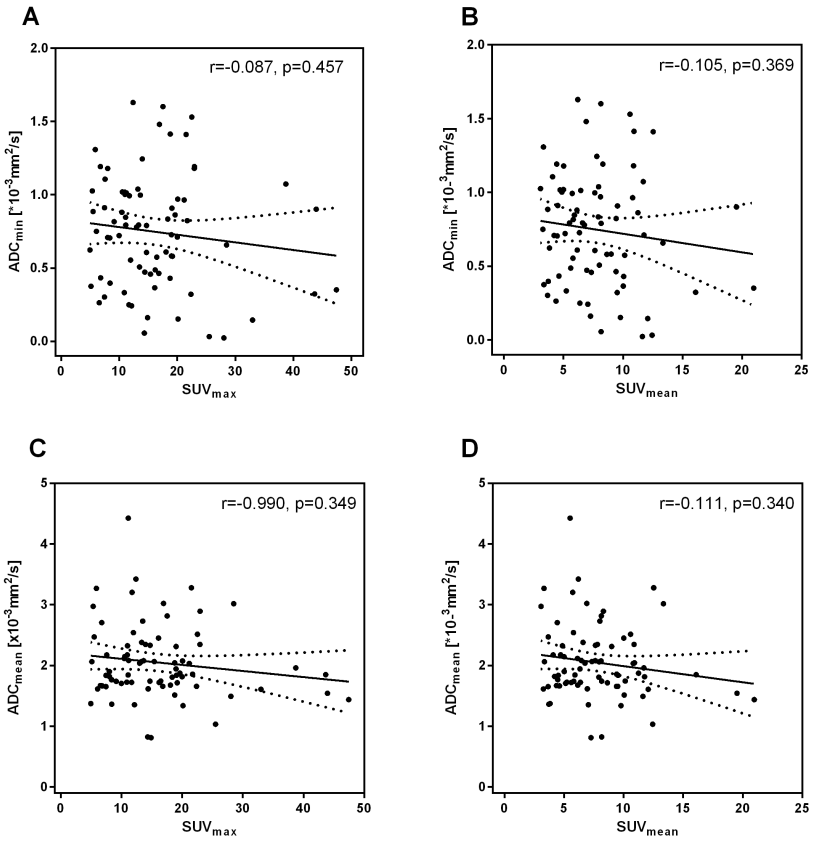


Figure 3.2: Scatter plots of correlation between tumor ADC and SUV values as determined on DW-MRI and ^{18}F-FDG-PET/CT imaging from 76 esophageal cancer tumors. For each scatterplot, the best-fit line is shown as the solid line. Dotted lines above and below represent the upper and lower 95% confidence intervals. **A**: ADC_{min} vs. SUV_{max}. **B**: ADC_{min} vs. SUV_{mean}. **C**: ADC_{mean} vs. SUV_{max}. **D**: ADC_{mean} vs. SUV_{mean}.

was 7.4 cm^3 (-26.9 - 41.7) (Figure 3.3). On average, the volumes determined on DW-MRI by semi-automatic delineation were larger compared to the volumes determined during semi-automatic contouring on ^{18}F-FDG-PET/CT.

Association of tumor ADC and SUV with clinical tumor characteristics

Table 3.2 shows the difference between functional imaging parameters (tumor ADC and SUV statistics) and established clinical tumor characteristics. In univariable analysis tumor ADC_{mean} and ADC_{min} showed no significant association with histologic tumor type, tumor grade, clinical T-stage and clinical N-stage. Although a trend towards higher SUV values in esophageal squamous cell carcinomas was observed, SUV_{mean} and SUV_{max} were not significantly associated with tumor characteristics.

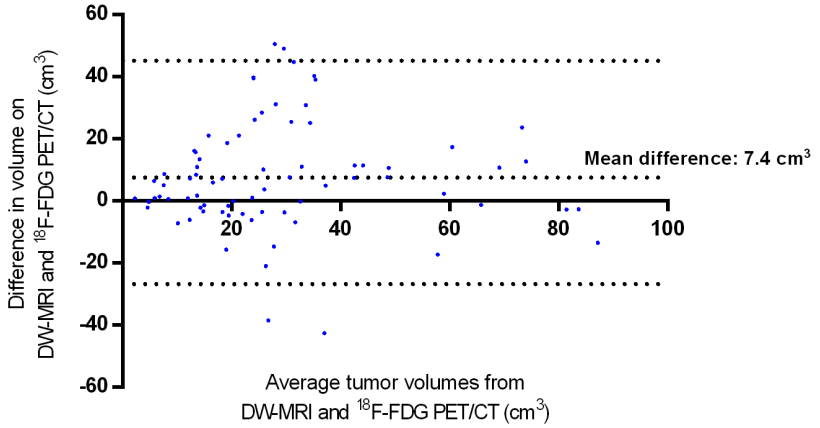


Figure 3.3: Bland-Altman plot comparing tumor volumes delineated on DW-MRI versus ^{18}F-FDG-PET/CT, respectively. The y-axis shows the difference between the two volumes (in cm^3), and the x-axis shows the mean of the two volumes. The black dotted lines represent the upper and lower 95% limits of agreement.

Discussion

In this multi-center prospective study, correlations between functional imaging markers derived from DW-MRI and ^{18}F-FDG-PET/CT imaging were evaluated. The results of the current study show that it was feasible to obtain complete data sets of both imaging modalities from most patients with newly diagnosed resectable esophageal cancer. No correlations were found between tumor ADC and SUV, regardless of the underlying histological subtype. Also, no associations of ADC and/or SUV with tumor stage or grade were found.

In order to assess potential clinical complementarity of tumor ADC and SUV values for prediction of survival or evaluation of response to treatment in esophageal cancer, the current study investigated the correlation between these two parameters. Assessment of treatment response is an important new determinant of prognosis, and guide towards more individualized treatment decisions for patients with locally advanced esophageal cancer [60, 102]. Due to the unsatisfactory results in treatment response assessment of currently available diagnostic techniques (i.e. endoscopic ultrasonography [103], endoscopy with biopsy [103], and ^{18}F-FDG-PET/CT [23, 24, 104], the utility of DW-MRI for response assessment in esophageal cancer is currently of high clinical interest. Although sequential ^{18}F-FDG-PET/CT scanning has shown to be moderately predictive for treatment response [23,24], ADC changes have shown great potential as marker for the degree of response to treatment in esophageal cancer [66]. The combination of these two functional imaging techniques may provide additional information with regard to tumor characterization.

Table 3.2: Association of ADC and SUV parameters with clinical tumor characteristics

Variable	n	Mean ADC (10^{-3}mm^2 s^{-1})	p	Lowest ADC (10^{-3}mm^2 s^{-1})	p	Mean SUV	p	Maximum SUV	p
Histologic tumor type*									
Adenocarcinoma	60	2.10 ± 0.6	0.289	0.75 ± 0.4	0.556	7.5 ± 3.5	0.068	15.7 ± 9.1	0.180
Squamous cell ca.	13	1.90 ± 0.3		0.82 ± 0.3		9.2 ± 3.6		18.3 ± 8.6	
Histologic tumor grade									
Moderate	43	1.97 ± 0.6	0.215	0.72 ± 0.4	0.535	7.4 ± 3.5	0.376	14.9 ± 8.0	0.247
Poor	33	2.15 ± 0.6		0.78 ± 0.4		8.1 ± 3.6		17.5 ± 9.9	
Clinical T-stage									
cT2	12	1.96 ± 0.6	0.567	0.60 ± 0.4	0.173	8.0 ± 3.3	0.574	16.7 ± 8.7	0.729
≥cT3	64	2.06 ± 0.6		0.77 ± 0.4		7.7 ± 3.6		15.9 ± 9.0	
Clinical N-stage									
cN0	25	1.90 ± 0.6	0.140	0.70 ± 0.4	0.471	7.5 ± 3.7	0.514	14.7 ± 8.9	0.216
cN+	51	2.12 ± 0.6		0.77 ± 0.4		7.9 ± 3.4		16.7 ± 8.9	

Data presented as mean ± standard deviation.
ADC: Apparent diffusion coefficient, SUV: Standardized uptake value.
*Patients with other histologic tumor types than adenocarcinoma and squamous cell carcinoma were excluded from this analysis.

Similar to our findings, previous studies assessing other types of cancer found no correlation between pretreatment tumor ADC and SUV values [105–109]. These results suggest that increased metabolic activity and restricted water diffusion represent independent phenomena and refer to different aspects of tumor pathophysiology. Combining multiparametric functional MRI with metabolic information derived from ^{18}F-FDG-PET/CT may therefore provide complementary information [106, 109].

Meanwhile, other studies comparing ADC and SUV in other types of cancer did observe a significant correlation between tumor ADC and SUV values [45, 110–115]. However, the reported correlation coefficients in some of these studies actually indicated a weak correlation (0-0.40) [112, 114, 115].

Several factors may account for the difference in literature with regard to reported correlations between tumor ADC and SUV. First, performing functional imaging in different types of cancer (e.g. ovarian and colorectal cancer lesions) results in significant different ADC and SUV values [109]. The heterogeneous outcomes of functional metrics in different types of cancer may partially account for this difference [113]. Second, several studies calculated the ADC_{mean} by taking averages of the mean ADC values of each separate image slice, regardless of the cross sectional size of the tumor on each slice [110–112]. In the current study the weighted average of all voxels of the entire VOI was taking into account, which results in a closer resemblance of the mean ADC of the whole tumor.

Tumor volume delineation was assessed on DW-MRI, and if necessary adapted after critical evaluation on the ADC maps. Although in general T_2-weighted images provide more anatomical details, this method ensured that the calculation of ADC was solely based on parameters obtained from high signal on DWI-MRI. A reasonable agreement of tumor volumes delineated on DW-MRI and ^{18}F-FDG-PET/CT was demonstrated in the Bland-Altman plot. On average, tumor volumes measured on DW-MRI were 7 cm^3 larger than those measured on ^{18}F-FDG-PET/CT. These findings correspond with the results of previous studies that found smaller tumor volumes on PET compared to DW-MRI imaging [110, 116, 117]. For the delineations on PET, a validated method with higher consistency compared to manual and thresholding methods was used [99]. However, MRI-based delineation in esophageal cancer tumors is relatively new. Currently, there is no consensus for accurate determination of tumor boundaries on MRI. Therefore, additional investigations aim to develop consensus on optimal MR delineation definitions for esophageal cancer.

Previous studies have reported associations between pre-treatment tumor ADC or SUV values, and histologic tumor characteristics (i.e. adenocarcinoma versus squamous cell carcinoma, tumor differentiation and cTN stage) in esophageal cancer [33, 66,118,119]. The current study found a trend towards higher SUV values in esophageal

squamous cell carcinomas, which is supported by previous studies [118, 119]. However, other associations could not be reproduced, which may be due to substantial differences in study population characteristics and sample size among the studies. Concerning ADC values, previous reports so far reported equivocal results, but were exploratory by nature and included a limited number of patients [66, 117].

Several limitations apply to this study. Factors potentially influencing the generalizability of the results include the hardware characteristics (i.e. different MRI and PET/CT systems), chosen imaging parameters and applied delineation techniques. Interobserver agreement analysis of ADC tumor delineation with a semi-automatic approach were not part of this study and need further elucidation. Moreover, the comparison between the imaging parameters (tumor ADC and SUV) and tumor characteristics (i.e. T-stage and histology) was based on small numbers which may have resulted in type 2 errors. Finally, pretreatment ADC, SUV and volume measurements could not be correlated to surgical pathology parameters, as the surgical specimens were obtained after neoadjuvant treatment in the current study. With respect to the aforementioned limitations, future comparative studies are needed to elucidate the advantages and disadvantages of DW-MRI and ^{18}F-FDG-PET/CT imaging in esophageal cancer. These studies may include comparison of DWI and PET parameters at multiple time points during treatment.

This study indicates that both tumor cellularity measured by DW-MRI imaging and tumor metabolism by ^{18}F-FDG-PET/CT are independent cellular phenomena in newly diagnosed esophageal cancer. Therefore, ADC and SUV values may have complementary roles as imaging markers in the prediction of survival and evaluation of response to treatment in esophageal cancer.

Chapter 4

DW-MRI and DCE-MRI are of complementary value in predicting pathologic response to neoadjuvant chemoradiotherapy for esophageal cancer

The following chapter is based on:

Sophie E. Heethuis, Lucas Goense, Peter S.N. van Rossum, A.S. Borggreve, S. Mook, Francine E.M. Voncken, Annemarieke Bartels-Rutten, Berthe M.P. Aleman, Richard van Hillegersberg, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier

Acta Oncologica (2018)

Abstract

Purpose
To explore the potential benefit and complementary value of a multiparametric approach using diffusion-weighted (DW-) and dynamic contrast-enhanced (DCE-) magnetic resonance imaging (MRI) for prediction of response to neoadjuvant chemoradiotherapy (nCRT) in esophageal cancer.

Material and Methods
Forty-five patients underwent both DW-MRI and DCE-MRI prior to nCRT (pre), during nCRT (week 2-3) (per) and after completion of nCRT, but prior to esophagectomy (post). Subsequently, histopathologic tumor regression grade (TRG) was assessed. Tumor apparent diffusion coefficient (ADC) and area-under-the-concentration time curve (AUC) were calculated for DW-MRI and DCE-MRI, respectively. The ability of these parameters to predict pathologic complete response (pCR, TRG1) or good response (GR, TRG$\leq$ 2) to nCRT was assessed. Furthermore the complementary value of DW-MRI and DCE-MRI was investigated.

Results
GR was found in 22 (49%) patients, of which 10 (22%) patients showed pCR. For DW-MRI, the 75^{th} percentile (P75) $\Delta\text{ADC}_{\text{post-pre}}$ was most predictive for GR (c-index = 0.75). For DCE-MRI, P90 $\Delta\text{AUC}_{\text{per-pre}}$ was most predictive for pCR (c-index = 0.79). Multivariable logistic regression analyses showed complementary value when combining DW-MRI and DCE-MRI for pCR prediction (c-index = 0.89).

Conclusions
Both DW-MRI and DCE-MRI are promising in predicting response to nCRT in esophageal cancer. Combining both modalities provides complementary information, resulting in a higher predictive value.

Introduction

Worldwide esophageal carcinoma continues to affect more than 450.000 people annually, with 5-year overall survival rates rarely exceeding 35% [2, 11, 12]. Currently, neoadjuvant chemoradiotherapy (nCRT) followed by surgery is considered standard of care for patients with resectable locally advanced esophageal cancer. Many studies have reported that the degree of response to neoadjuvant therapy is associated with patient prognosis, with pathologic complete response (pCR) having the most favorable long-term prognosis [11, 14, 58, 60, 61]. Accurate prediction of good response to nCRT potentially allows a wait-and-see approach, with omission of surgery and close clinical follow-up. On the other hand, accurate prediction of poor response allows early treatment modification, e.g. intensification of neoadjuvant therapy or primary surgery.

Unfortunately, diagnostic modalities including endoscopic biopsy, endoscopic ultrasonography and computed tomography (CT) that are currently used to identify pathologic response yield unsatisfactory results [103,120]. Highest overall pooled sensitivities and specificities were found for ^{18}F-FDG-PET(/CT), however, with values ranging from 67% to 70% this is still insufficient to justify changes in clinical decision-making [23, 24]. Meanwhile, functional MRI techniques are emerging as an advanced imaging technique since it visualizes different and possibly complementary tumor characteristics. With diffusion-weighted (DW-) magnetic resonance imaging (MRI) the variation in free diffusion of water molecules between different tissues is visualized, allowing it to distinguish tumorous tissue with high cellular density from healthy tissue using the apparent diffusion coefficient (ADC) [37, 38, 121]. With dynamic contrast-enhanced (DCE-) MRI, the vascular integrity is visualized using administration of a contrast agent, allowing to monitor neoplastic changes in tissue [122]. Two recent pilot studies reported that tumor alteration on both DW-MRI and DCE-MR imaging during the first 2-3 weeks of nCRT for esophageal cancer are highly predictive for pCR [66,123]. However, these findings have not yet been validated in a larger cohort. Also, no direct comparison between DW-MRI and DCE-MRI for treatment response evaluation in this setting has been reported.

The purpose of the current study was to evaluate, in a multicenter setting, whether combining data from both DW-MRI and DCE-MRI in patients receiving nCRT for esophageal cancer yields complementary information for response prediction, and may therefore increase predictive power compared to single technique MR imaging.

Table 4.1: Patient and treatment-related characteristics for both institutes.

	Study population UMCU (n=31)	Study population NKI (n=14)	Combined analysis DWI/DCE UMCU+NKI (n=34)
Gender			
Male	27	11	29
Female	4	3	5
Age, years (at start RT)*	64.5 ± 7.6	58.6 ± 13.0	62.4 ± 10.1
Clinical T-stage			
T2	6 (19%)	2 (14%)	5 (15%)
T3	25 (81%)	12 (86%)	29 (85%)
Clinical N-stage			
N0	9 (29%)	6 (43%)	10 (29%)
N1	8 (26%)	5 (36%)	12 (35%)
N2	12 (39%)	3 (21%)	9 (26%)
N3	2 (6%)	0 (0%)	3 (9%)
Type			
SCC	3 (10%)	2 (14%)	4 (12%)
AC	26 (84%)	12 (86%)	28 (82%)
ASC	2 (6%)	0 (0%)	2 (6%)
Location			
Proximal third of esophagus	1 (3%)	0 (0%)	1 (3%)
Middle third of esophagus	3 (10%)	1 (7%)	3 (9%)
Distal third of esophagus	23 (74%)	6 (43%)	21 (62%)
Gastroesophageal junction	4 (13%)	7 (50%)	9 (26%)
Acquisition, no. of days			
Pre (before start nCRT) †	6 (2 - 14)	5 (-1 - 11)	6 (-1 - 14)
Per (after start nCRT)†	10 (8 - 17)	15 (9 - 16)	10 (8 - 17)
Post (after completion nCRT)†	41 (17 - 62)	39 (31 - 65)	40.5 (17 - 57)
Post scan - surgery interval†	10 (5 - 48)	11.5 (4 - 19)	11.5 (4 - 48)

The patient characteristics from the combined analysis are summarized in the third column. SCC = squamous cell carcinoma; AC = adenocarcinoma ; ASC = adenosquamous carcinoma. Data presented as counts with percentages in the parentheses unless stated otherwise.
* Data presented as mean ± SD.
† Data presented as median (range).

Material and methods

Study population

Patients were included in the University Medical Center Utrecht (UMCU) and in the Netherlands Cancer Institute (NKI), from August 2013 to January 2016. Written

informed consent was obtained from all patients and this study was approved by the institutional review board of each center. Patients with contraindications for 1.5 T MRI with administration of a contrast agent were not eligible for inclusion. All 46 patients with biopsy-proven esophageal cancer receiving nCRT followed by surgery who underwent all MRI studies were initially included in this study. One patient was excluded due to a very small tumor volume (<2.5 mL on the initial scan). Of the remaining patients, 31 were scanned in the UMCU and 14 in the NKI (Table 4.1).

Treatment protocol

All patients included in this study were planned to receive weekly administration of intravenous carboplatin (area under the curve of 2 mg/mL per minute) and paclitaxel $50\,\mathrm{mg\,m^{-2}}$ (50 mg/m^2 body-surface area) for 5 weeks, with concurrent radiotherapy of 41.4 Gy in 23 fractions of 1.8 Gy [11]. All patients were treated with step-and-shoot intensity modulated radiation therapy (IMRT) or volumetric modulated arc radiotherapy (VMAT), planning goals were to cover at least 99% of the target volume with a minimum of 95% of the prescribed dose. Five to ten weeks (median 8 weeks) after completion of neoadjuvant treatment, all patients underwent a transhiatal or transthoracic esophagectomy.

Histopathologic assessment

Histopathologic assessment of the resection specimen was performed by a dedicated pathologist with GI subspecialty according to Mandard, with a tumor regression grade (TRG) ranging from 1 (pCR) to 5 (absence of regressive changes) [17]. Two approaches were applied to discriminate between response to treatment: an analysis was performed to differentiate between patients with good (GR, TRG$\leq$2) versus poor response (noGR, TRG$\geq$3), and pathologic complete response (pCR, TRG1) versus no pathologic complete response (no-pCR, TRG$\geq$2).

MRI acquisition

Acquisition of MR images was performed at three time points: a baseline scan prior to treatment (pre), a scan after two weeks (8-13 fractions) of nCRT treatment (per) and a preoperative scan, 3-9 weeks after completion of nCRT treatment (post). In all patients there was a similar initial intention to schedule surgery 6-8 weeks after completion of nCRT and plan the post-scan within two weeks prior to surgery. All image acquisition was performed on a 1.5 T Philips Achieva or Ingenia (Best, The Netherlands) using the Torso (16 channels) or Anterior/Posterior (28 channels) receive coils, respectively. The same imaging protocol was used in both institutes. DWI scans were acquired in transverse plane during free breathing, for three different b-values (0, 200, 800 $\mathrm{s\,mm^{-2}}$), with a bandwidth per pixel of 29.4 or 23.4 Hz for the Achieva and Ingenia, respectively. The DCE-series consisted of 62 3D scans, scanned with a

Table 4.2: Scan parameters of the used MRI protocol

	DWI EPI (STIR)	DCE (3DFFE T_1W)	T_2W (MS-TSE)	B_0 †
Scan plane	transverse	coronal	transverse	transverse
Slice thickness (mm)	4.00	3.00	4.00 (gap: 2.48)	4.00
Voxel size (mm)	3.25	1.18 or 1.98*	0.67	4.06
TR (ms)	7241	3.43	1604	630.4
TE (ms) (TE_2)	76.2	1.53	100	4.6 (9.2)
Flip angle (°)	90	20	90	30
Temporal resolution (sec)	*N.A.*	3	*N.A.*	*N.A.*
Number of time frames	*N.A.*	62	*N.A.*	*N.A.*
b-values ($s\,mm^{-2}$)	0, 200, 800	*N.A.*	*N.A.*	*N.A.*

N.A. = not applicable.
† Dual acquisition ; * second parameter applicable for scans acquired with Ingenia scanner.

temporal resolution of 3 seconds. After the 10^{th} scan the contrast agent was injected with an automatic syringe pump at a flow rate of 1 mL/sec (for patients >100 kg 2 mL/sec), followed by a saline injection. In the UMCU gadobutrol (Gd-BT-DO3A, Gadovist; Schering AG, Berlin, Germany) was used at a dose of 0.1 mmol/kg of body weight, at the NKI Dotarem (Gadoteric acid, 0.5 mM; Guerbet, Paris, France) was used with a fixed dose of 7.5 mmol for each patient. In order to prevent artifacts in the aorta due to pulsatile flow, the heart was included in the scanned volume. Prior to the DCE scan, a transverse T_2-weighted scan was acquired for anatomical verification with a multi-slice turbo spin echo sequence, using a navigator for respiratory triggering [66, 123]. Detailed scanning parameters are presented in Table 4.2.

Image analysis

DWI scans were corrected for geometric distortions using a B_0-field inhomogeneity map [124] and ADC values were calculated using a mono-exponential model. DCE scans were processed according to a previously reported protocol [123]. Rigid registration was performed to account for motion between scans. A series of scans with variable flip-angles were acquired. This enabled quantification of the intrinsic tissue T_1 time which is required to quantify the contrast agent concentration after injection [125]. The area-under-the-concentration versus time curve (AUC), defined as an integral over 60 seconds after inflow of contrast, was calculated. Examples of a T_2-weighted, a DWI and a DCE scan and the corresponding ADC- and AUC-maps are shown in Figure 4.1. To minimize inter-observer variability, the primary tumor was automatically delineated on the initial b800 DWI using a contouring algorithm in ITK-SNAP v3.4.0.QT4 with 2 clusters in pre-segmentation mode and default evolution parameters [98]. As a result, a region with high values on the b800 DWI was delineated. These automatic delineations were verified and manually adjusted if

necessary (e.g. in case of volume overestimation or the inclusion of lymph nodes), using both the ADC-map and T_2-weighted scan in Volumetool, an in-house developed image software tool [126]. This pre-delineation was registered to the next time point (per) and manually adjusted for possible shrinkage in the circumference of the esophagus, which was then propagated to the final time point (post), again followed by manual adjustments if necessary. In other words, in each individual patient the cranio-caudal length of the delineations remained constant over the three time points. All adjustments were performed in consensus by two readers (S.E.H. and L.G.) and verified by a radiation oncologist (S.M.).

Data analysis

In line with our previous work [66, 123], mean, median and several percentiles (P75/P90) were calculated for ADC (DW-MRI) and AUC (DCE-MRI) within the delineations on both modalities. Separate time points as well as percentage differences between time points were analyzed. The first acquired scan was used as a reference for these calculations, as this was found to provide the highest predictive performance in the prior studies on DW-MRI or DCE-MRI. Previously published optimal thresholds for both modalities were reoptimized in this larger patient cohort.

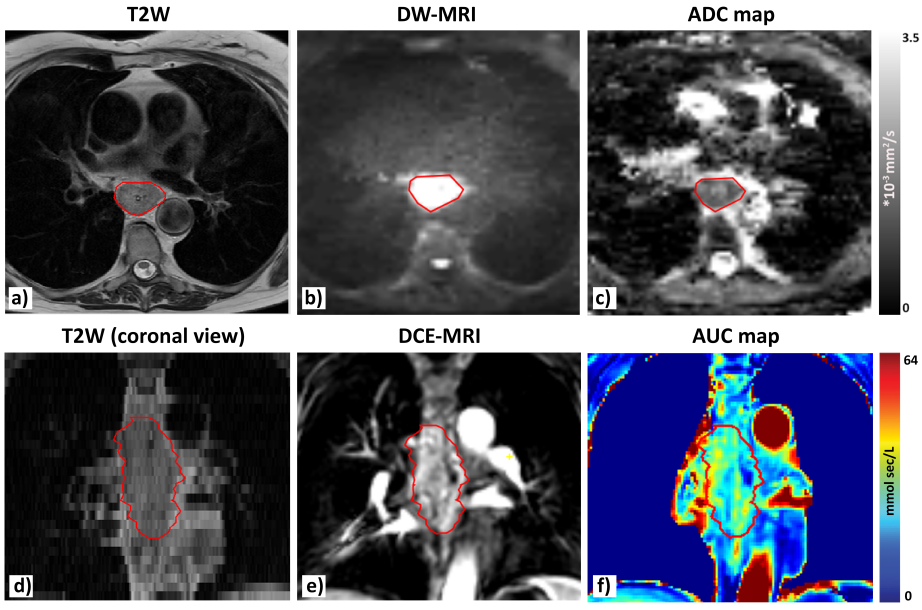


Figure 4.1: An example of acquired MRI scans for one patient for one time point (pre scan). In **(a)** a T_2W scan is shown with in **(d)** the coronal view of the same scan. **(b)** A b800 DW-MRI scan and its corresponding ADC-map in **(c)**. In **(e)** the 24th scan of a DCE-MRI series is shown, during inflow of contrast in the tumor. The corresponding calculated AUC-map is found in **(f)**.

A quantitative quality assessment was performed on all DW-MRI scans because some scans were visually of poor quality. Scans with poor quality typically showed high noise levels in the lung and/or poor visibility of the spleen, therefore on the b800 DWI both organs were delineated and mean signal values were divided to reach a quantitative quality measure (spleen/lung). The analysis of the DW-MRI and a combined analysis of DW-MRI and DCE-MRI was only performed in scans in which the aforementioned quality measure was higher than mean minus standard deviation (SD) of the whole patient group.

Statistical analysis

Patient and treatment-related characteristics were described as count with percentages, mean with SD or median with range as appropriate. The association between response and DW-MRI or DCE-MRI parameters was estimated using univariable logistic regression analysis. Diagnostic performance measures (i.e. sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)) were computed and receiver operating characteristics (ROC) curve analysis was performed, calculating the area-under-the-curve (c-index). A p-value of <0.05 was considered statistically significant. Complementary value of DW-MRI and DCE-MRI was assessed using a multivariable logistic regression model. Model performance was assessed by the Akaike Information criterion [127]. Predictive values were only presented when DW-MRI and DCE-MRI were complementary. Ideal cut-off values were calculated by giving equal weight to sensitivity and specificity. The threshold in terms of probability (output of logistic regression) were converted in terms in of ADC and AUC following the equation:

$$\text{Probability (pr)} = \frac{e^{(\alpha+\beta_1 \cdot X_1+\beta_2 \cdot X_2+..+\beta_n \cdot X_n)}}{1+e^{(\alpha+\beta_1 \cdot X_1+\beta_2 \cdot X_2+..+\beta_n \cdot X_n)}} \tag{4.1}$$

with n the number of variables X (univariable [n=1], techniques combined [n=2]), and α/β constants calculated by the logistic regression model. All statistical analyses were performed in SPSS Statistics version 22 (IBM Corp., Armonk, NY, USA) and visual representations using GraphPad Prism 6.07 software (GraphPad Software, La Jolla, CA, USA).

Results

For the total group of 45 patients, assessment of the histopathologic tumor type revealed adenocarcinoma in 38 (84%) patients, squamous cell carcinoma in 5 (11%) patients, and adenosquamous carcinoma in 2 (4%) patients. Patient- and treatment-related characteristics are shown in Table 4.1. The resection specimen showed good response (TRG 1-2) in 22 (49%) patients, of which 10 (22%) patients showed pCR.

Table 4.3: Predictive values for both DW-MRI and DCE-MRI discriminating pCR versus no-pCR and GR versus noGR.

	n	Threshold	Sens. (%)	Spec. (%)	PPV (%)	NPV (%)	c-index	p-value
GR vs. noGR								
P90 $\Delta AUC_{per\text{-}pre}$	45	33.7%	68	57	60	65	0.64	0.261
P75 $\Delta ADC_{post\text{-}pre}$	34	22.4%	87	58	62	85	0.75	0.031
pCR vs. no-pCR								
P90 $\Delta AUC_{per\text{-}pre}$	45	27.7%	90	63	41	96	0.79	0.028
P75 $\Delta ADC_{post\text{-}pre}$	34	29.5%	83	54	28	94	0.76	0.050
Combined	34	0.138*	100	75	46	100	0.89	0.086/0.052**

Only the most predictive parameters are shown. Also a combined analysis using multivariable logistic regression, which was found to be complementary, is presented.
Sens. = Sensitivity ; Spec. = Specificity
* Probability value following Equation 1 with model values: α=-2.244, β_{AUC}=-0.034 and β_{ADC}=0.044.
** for P90 $\Delta AUC_{per\text{-}pre}$ and P75 $\Delta ADC_{post\text{-}pre}$ respectively.

DCE-MRI parameter AUC_{per} 90th percentile (P90) showed a significant difference in separating pCR from no-pCR (p = 0.044, c-index = 0.72). An initial increase followed by a decrease was found when comparing the three time points. AUC P90 showed for pCR values of 34.0 ± 9.4 mmol L^{-1} s [mean ± SD], 36.8 ± 8.3 mmol L^{-1} s and 29.5 ± 13.9 mmol L^{-1} s for pre, per and post, respectively. Similarly, AUC P90 for no-pCR was 34.3 ± 9.5 mmol L^{-1} s, 48.9 ± 16.1 mmol L^{-1} s and 32.0 ± 9.6 mmol L^{-1} s, for pre, per and post, respectively. When comparing differences between scans with respect to the first time point, the most predictive value to separate pCR from no-pCR, was the relative increase in tumor AUC with a P90 $\Delta AUC_{per\text{-}pre}$ of 10.6% ± 17.6% and 45.2% ± 41.5% for pCR and no-pCR, respectively (p = 0.028, c-index = 0.79) (Figure 4.2a,b). No significant differences were found for ΔAUC in predicting GR versus noGR (Table 4.3).

Quantitative quality assessment of the b800 maps for the three different time points resulted in spleen/lung image intensity ratio [mean ± SD] values of 5.9 ± 2.0, 5.0 ± 1.7 and 4.4 ± 1.7 for pre, per and post scans, respectively. A total of 24 out of 138 (17%) scans revealed a value below mean-SD, with a large overlap within the same patients for the three time points. This resulted in exclusion of 11 patients (AVL [n=5] and UMCU [n=6]), of which 7 GR and 4 noGR (Table 4.1, 3rd column). No direct relation was found between excluded patients and patient and treatment-related characteristics (e.g. TN-stage, tumor location). For the remaining 34 patients, analysis was performed on DW-MRI parameter ADC. Analysis of the three time points showed only a significant difference in separating GR from noGR for ADC_{per} P90 (p = 0.040, c-index = 0.70) and ADC_{post} P90 (p = 0.040, c-index = 0.70). An increasing trend in ADC over time was found for both good and poor responders (median ADC for GR 1.87 ± 0.41 · 10^{-3} mm^2 s^{-1} [mean ± SD], 2.31 ± 0.37 · 10^{-3} mm^2 s^{-1} and

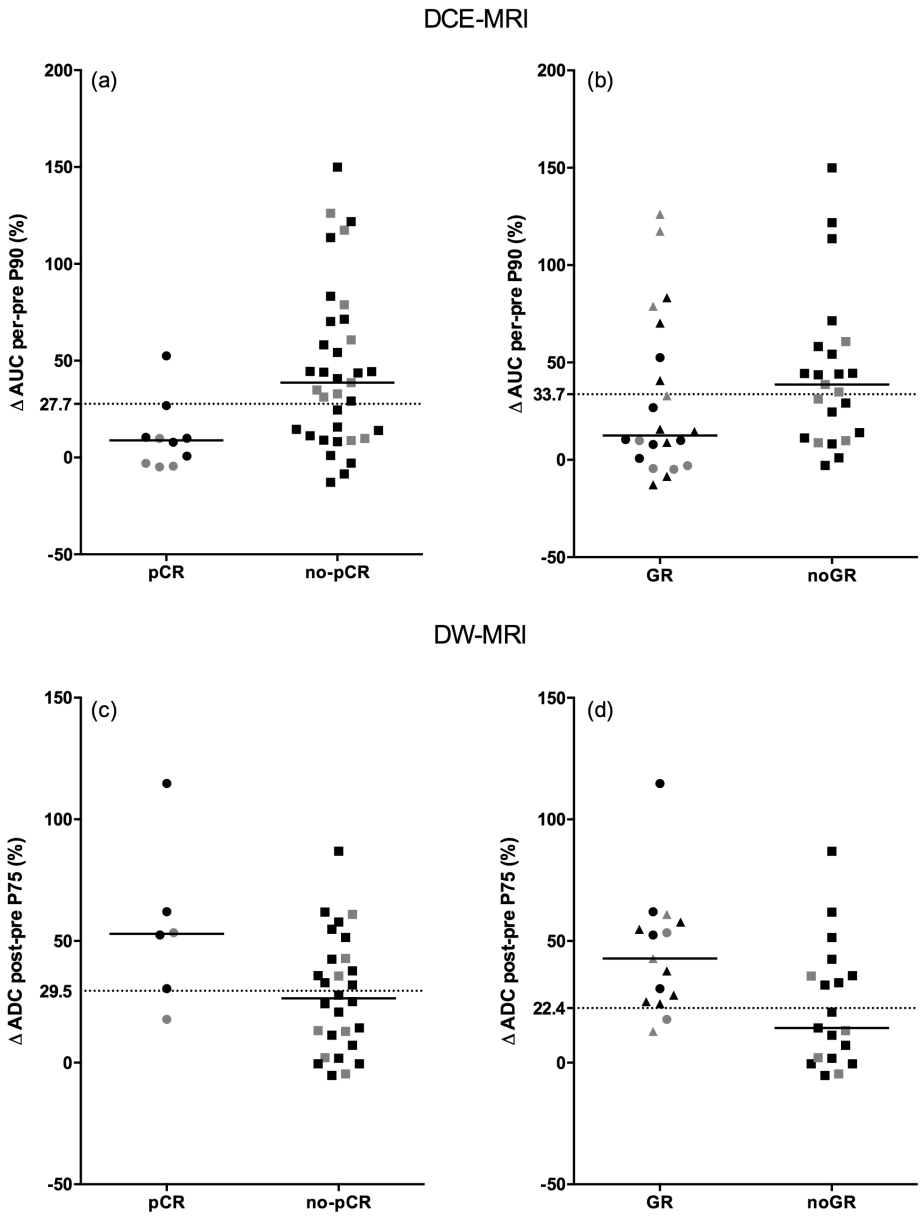


Figure 4.2: In **(a)** and **(b)** 90th percentile (P90) $\Delta AUC_{per\text{-}pre}$ is presented as a function of response and in **(c)** and **(d)** 75th percentile (P75) $\Delta ADC_{post\text{-}pre}$. In **(a,c)** pathologic complete response (pCR) and no-pCR are differentiated while in **(b,d)** good response (GR) versus poor response is shown. Each dot represents a patient and the two institutes are indicated separately (UMCU in black/NKI in grey). The solid black line for each group represents the median. Threshold lines are indicated for which predictive values are calculated for separation based on response.

$2.72 \pm 0.62 \cdot 10^{-3}$ mm^2 s^{-1} for pre, per and post, respectively, and for noGR $1.96 \pm 0.32 \cdot 10^{-3}$ mm^2 s^{-1}, $2.15 \pm 0.37 \cdot 10^{-3}$ mm^2 s^{-1} and $2.41 \pm 0.51 \cdot 10^{-3}$ mm^2 s^{-1}). Significant associations with response were found in ΔADC with respect to the first time point. $\Delta ADC_{per\text{-}pre}$ P90 resulted in a [mean ± SD] of 23.5% ± 20.5% and 9.8% ± 11.7% for GR and noGR respectively (p = 0.035, c-index = 0.70). The most predictive parameter was $\Delta ADC_{post\text{-}pre}$ P75 with [mean ± SD] of 45.0% ± 25.3% and 23.1% ± 25.2% for GR and noGR respectively (p = 0.031, c-index = 0.75) (Figure 4.2c,d). Analysis of ΔADC in predicting pCR versus no-pCR did not show significant associations (Table 4.3).

Multivariable logistic regression analysis, including patients for whom both DW-MRI and DCE-MRI were available (n=34), revealed that these parameters were complementary in predicting pCR (c-index = 0.89). Predictive values of this model and the univariable analysis are presented in Table 4.3. The $\Delta AUC_{per\text{-}pre}$ as function of $\Delta ADC_{post\text{-}pre}$ is presented in Figure 4.3 in which threshold lines of both univariable models as well as the multivariable model are shown. The ROC-curves of the logistic regression models are shown in Supplementary figure 4.1a for both GR and pCR, in which the chosen threshold for the complementary model predicting pCR is marked. Since percentile values are not commonly produced in analysis software, predictive values and ROC-curves are also presented for mean $\Delta AUC_{per\text{-}pre}$ and $\Delta ADC_{post\text{-}pre}$ in Supplementary Table 4.1 and Supplementary Figure 4.1b. Slightly lower c-index and predictive values were reached, multivariable logistic regression showed no complementary value.

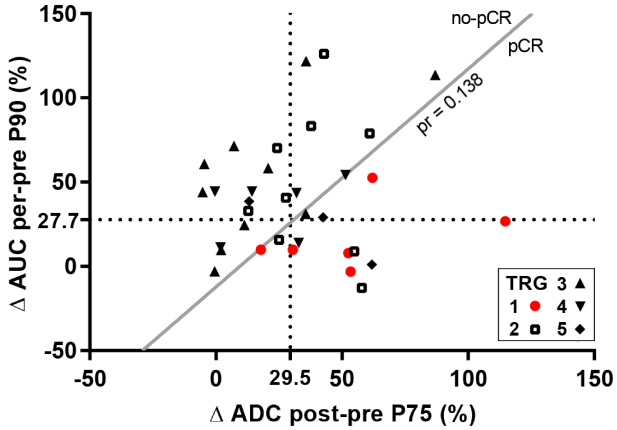


Figure 4.3: 90th percentile (P90) $\Delta AUC_{per\text{-}pre}$ is plotted as function of 75th percentile (P75) $\Delta ADC_{post\text{-}pre}$. Each dot represents a patient and TRG scores are differentiated using different symbols, with pathologic complete response (pCR) highlighted in red. Equal thresholds from Fig. 4.2 predicting pCR are indicated in dotted lines. The solid line represents the threshold line (separating pCR from no-pCR) from the multivariable logistic regression model, for which the probability equals 0.138.

Discussion

This multi-center prospective study demonstrates that DW-MRI and DCE-MRI provide complementary information for predicting response to nCRT in patients with esophageal cancer. Separate analysis of both DW-MRI and DCE-MRI showed trends in discriminating good from poor responders comparable to previously published work [66, 123].

To our knowledge, this is the first study investigating the complementary value of DW-MRI and DCE-MRI for response prediction in patients with esophageal cancer. A previous published study assessed the independent value of intravoxel incoherent motion (IVIM) MRI and DCE-MRI in patients with esophageal cancer, but did not report a combined analysis [128]. In other tumor sites multiple studies have reported superior descriptive accuracies of combined analyses compared to separate analyses [42, 43].

Both DW-MRI and DCE-MRI showed significant associations with response when ADC and AUC were analyzed at separate time points. However, higher c-index values and significant associations with response were found for both modalities when differences between time points were analyzed, which is in correspondence with our previously published work and publications concerning other tumor sites [44, 66, 123, 129]. In the DCE-MRI series, the most predictive parameter was the difference between the pre-scan and the scan acquired during nCRT. Analyses of DW-MRI series showed, in contrast to previous published work [66], that differences between the post and baseline-scan had a higher predictive value than differences between baseline and per-scans. Correct prediction of response during treatment may enable early modification of the treatment (i.e. nCRT intensification or discontinuation) while prediction of response using a post-treatment scan can guide the choice to omit surgery and possibly use a wait-and-see approach [130, 131].
Combining both modalities increased predictive values and showed that they complement each other (Fig. 4.3). In this study, uni- and multivariable logistic regression models were used to predict response to nCRT in patients with esophageal cancer. To further optimize the prediction model, the addition of other modalities could be examined, as well as the use of other models (e.g. machine-learning methods) [132, 133].

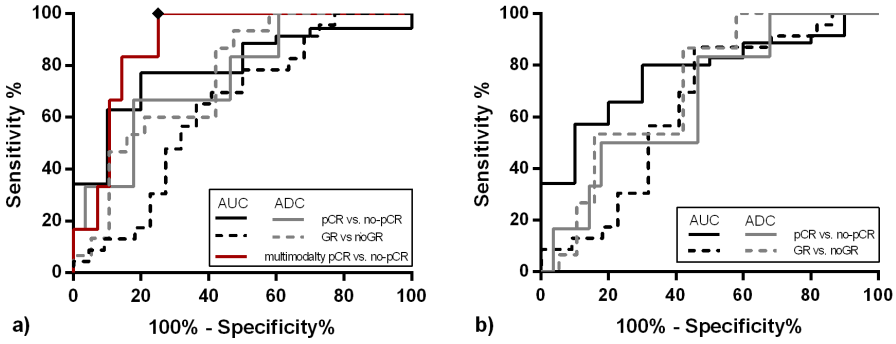
A different approach was chosen for delineation of the tumor compared to our previous work. To minimize inter-observer variability, the pre-scan was delineated using automatic-contouring software [98], which was then transformed to the following time point. Adjustments for possible anatomical changes in esophageal circumference were made using both the ADC-map and anatomical T_2-weighted images. This method is less user-dependent and more sensitive to detect quantitative differences within the tumor over time, because it includes the cranio-caudal length and circumference of the initial tumor on all MR scans. The delineation is thus not solely based on the

high b800 signal intensity of each time point separately. This could explain why the $\Delta ADC_{post\text{-}pre}$ was found to be more sensitive to response prediction compared to our previous study [134]. Furthermore, in the current study the DCE-MRI parameter AUC resulted in higher predictive values compared to our previous results, what may be caused by the differences in delineation approaches between this study and our previous smaller DCE-study [123].

In the determination of a threshold for separating responders from non-responders, a trade-off is made between a preference for high specificity (proportion of poor responders well classified), high sensitivity (proportion of good responders well classified), high PPV or high NPV, depending on the purpose. The threshold can be changed to achieve for example a high NPV when the goal is to justify therapy discontinuation in poor responders.

Some limitations apply to this study. First, the reproducibility within patients of the researched parameters is unknown. However, the ADC calculation was validated using a diffusion phantom in both institutes (HPD, Boulder, USA). Second, the exclusion of patients due to poor scan quality resulted in a relatively small sample size of the DW-MRI and multivariable logistic regression analyses, therefore leading to limited clinical impact. Although poor scan quality was found to be patient dependent (i.e. the same patients showed poor scan quality at multiple time points), no relations were found between excluded patients and clinical parameters. After the current study, adjustments to the imaging protocol were made. Promising results were obtained with higher and more stable scan quality and less influence of the presence of breathing motion, which could have possibly influenced the predictive values found in this study. A multicenter prospective validation study in a large cohort of patients is currently underway, which should resolve the limitations in the current literature (ClinicalTrials.gov identifier NCT03474341). The current study justifies studying both modalities, since they showed to be of complementary value. Third, the timing of the second scan was based on our previously published work [24,66,123]. However, a study that is currently being performed in our institute is assessing the optimal time point for response prediction during the course of nCRT, which could further improve the reported predictive values. The timing of surgery was intended to be 6-8 weeks after nCRT with a post-scan within two weeks prior to surgery. Due to for example logistic reasons this date varied. Variations in time between neoadjuvant therapy and surgery impact the observed pathologic response and could therefore influence our results [135, 136].

In conclusion, our multicenter study shows that changes in both DW-MRI and DCE-MRI parameters during treatment are promising imaging markers for prediction of response to nCRT in patients with esophageal cancer. Furthermore, both modalities provide distinct predictive information about the tumor, resulting in increased accuracy when using a multiparametric response prediction model.



Supplementary Figure 4.1: (a) ROC-curve for univariable DW-MRI and DCE-MRI percentile values (P90/P75) as presented in Table 4.3. For pCR versus no-pCR the ROC-curve of the complementary multimodality analysis is shown in red with the diamond marker indicating the chosen threshold. This point can be varied depending on the desired weight towards high sensitivity, specificity, PPV or NPV. **(b)** ROC-curves for the supplementary results using only mean-values for both AUC and ADC.

Supplementary Table 4.1: Predictive values for both DW-MRI and DCE-MRI discriminating pCR versus no-pCR and GR versus no-GR using mean values for both modalities. Indicated p-values are calculated using univariable logistic regression.

	n	Threshold	Sens. (%)	Spec. (%)	PPV (%)	NPV (%)	c-index	p-value
GR vs. noGR								
mean $\Delta AUC_{per\text{-}pre}$	45	29.0	68	57	60	65	0.64	0.248
mean $\Delta ADC_{post\text{-}pre}$	34	19.6	87	58	62	85	0.72	0.067
pCR vs. no-pCR								
mean $\Delta AUC_{per\text{-}pre}$	45	27.6	90	57	38	95	0.77	0.036
mean $\Delta ADC_{post\text{-}pre}$	34	25.4	83	54	28	94	0.67	0.190

Sens. = Sensitivity ; Spec. = Specificity.

Chapter 5

Analysis of intravoxel incoherent motion (IVIM) magnetic resonance imaging data for esophageal cancer using different fitting approaches of a bi-exponential model

The following chapter is based on:

Sophie E. Heethuis, A.S. Borggreve, Lucas Goense, Peter S.N. van Rossum, S. Mook, Richard van Hillegersberg, Jelle P.Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier

Submitted

Abstract

Purpose
The aim of this methodological study was to assess the impact of three fitting approaches on the resulting intravoxel incoherent motion (IVIM) MRI parameters in patients with esophageal cancer.

Method
Twenty-four patients treated with neoadjuvant chemoradiotherapy (nCRT) for esophageal cancer underwent six sequential MRI scans. Scans were acquired prior to the start of nCRT, followed by weekly MRI scans during nCRT. A coronal IVIM-MRI was acquired with 13 b-values ranging from 0 to 800 $\mathrm{s\,mm^{-2}}$. The performance of three different fitting approaches to the Bihan bi-exponential function computing IVIM-MRI-related diffusion and perfusion parameters was compared at voxel-level within the suspected tumor region. For reference purposes two conventional mono-exponential models were fitted using three b-values.

Results
A total of 47.294 tumor voxels were analyzed by three different fitting approaches to the bi-exponential Bihan function using Bland-Altman density plots. All three fitting approaches resulted in good fits between data and model (median R^2 values between 0.96 and 0.97). However the fitted model parameters between the approaches varied, which was especially apparent in the pseudo-diffusion D^* with median values of 22.0, 28.9 and 32.7 $\cdot 10^{-3}$ $\mathrm{mm^2\,s^{-1}}$ for the Bihan-fixed, Bihan-free and Bayesian approaches, respectively. Despite the high R^2 values, the Bihan-free approach showed large outliers, especially for D^*.

Conclusion
Although only small differences were found in fitting performance, large discrepancies within the model parameters were observed between various fitting approaches for IVIM-MRI data. It is expected that these differences influence the accuracy and precision in longitudinal IVIM studies for response assessment. Therefore, we strongly argue to explicitly report specifications of fitting approaches used in future IVIM studies to be able to compare outcomes between studies.

Introduction

Intravoxel incoherent motion (IVIM) magnetic resonance imaging (MRI) is increasingly investigated in relation to modeling tumor characteristics. In diffusion-weighted (DW) MRI typically two to three scans are acquired with different diffusion weights (b-values), followed by calculation of the apparent diffusion coefficient (ADC) using a mono-exponential function. Many studies research the potential of tumor ADC values and their relation to treatment outcome [37, 38, 44, 66]. However, already in the late 80's Bihan et al. demonstrated that the signal decay as a function of b-values is not mono-exponential and depends on perfusion in the lower range of b-values [39]. To account for this phenomenon, a two-compartment model was proposed. This model more accurately reflects tumor biology and heterogeneity, and as a consequence could have large potential in assessing treatment-induced changes.

Esophageal cancer is known to have a poor 5-year overall survival rate. The 29% of patients who have a pathologic complete response after neoadjuvant chemoradiotherapy (nCRT) followed by an esophagectomy have the most favorable long-term prognosis [11, 12, 58, 60, 61]. Since accurate prediction of response potentially allow for omission of esophagectomy, multiple studies have assessed the potential of several modalities to predict response to nCRT in esophageal cancer [66, 120, 123, 137]. Both DW-MRI (diffusion) and dynamic-contrast enhanced (DCE) MRI (perfusion) have shown to serve as good predictors for response [66, 123]. As IVIM theoretically combines information on both perfusion and diffusion, it is of interest to investigate its usefulness for response assessment in esophageal cancer. The use of IVIM instead of DCE (combined with DWI) would spare the patient from intravenous contrast injection [128, 138].

Before investigating the correlation between IVIM and other modalities, establishing a reliable fitting procedure is of great importance. Multiple studies were published investigating different models and fitting approaches [40,139–141], demonstrating that substantial discrepancies can occur in precision and accuracy depending on the applied fitting approach [40]. However, the voxel-based influence of using different approaches for IVIM-MRI fitting in esophageal cancer is unclear yet. In view of current and future large multi-center trials for response assessment in patients with esophageal cancer including IVIM-MRI, the aim of this methodological study was to assess the influence of using three different fitting approaches. For this purpose an extensive IVIM-MRI dataset of patients with esophageal cancer, acquired at six time points per patient, during the course of nCRT was analyzed.

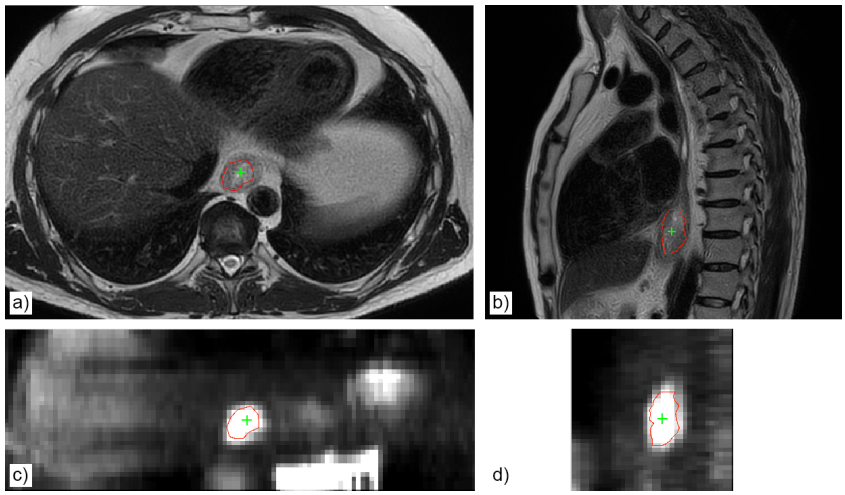


Figure 5.1: The transverse **(a)** and sagittal **(b)** anatomical T_2-weighted MultiVane scan of the baseline scan of a patient with adenocarcinoma. In **(c)** and **(d)** the transverse and sagittal view (zoomed in) of the b800 scan, linked to the T_2-weighted scans (see green crosshair). It is clearly visible that the tumor shows a hyperintense signal on the IVIM scan. The tumor delineation is depicted in red in all images. It should be noted that this delineation does not represent a clinical delineation of the tumor, but represent a slightly reduced volume for IVIM analysis purposes.

Methods and Materials

Study population

This prospective observational single-center study was approved by the local medical ethics committee and written informed consent was obtained from all patients. Patients with contraindications for MRI were not eligible for inclusion. A total of 27 consecutive patients were included, between March 2016 and October 2017. All patients were scheduled to undergo esophagectomy, followed by pathologic assessment of treatment response. Three patients were excluded from analysis due to poor visibility of the tumor on diffusion-weighted images (n=2) and withdrawal of informed consent (n=1). Of the remaining 24 patients included for analysis, in two patients one of the six time points was missing, in one patient a scan session was stopped before the IVIM series was scanned and in one patient a scan session suffered from poor tumor visibility, resulting in a total of 140 scan series available for analysis.

MRI acquisition

MR images were acquired at six time points; once prior to the start of treatment followed by five weekly MR scans during the course of nCRT. All images were acquired on a 1.5 T MRI system (Philips Ingenia, Best, the Netherlands), using the Anterior/Posterior (28 channel) receive coils. A respiratory-triggered anatomical T_2-weighted scan was acquired with a multi-slice turbo spin echo sequence (TR/TE =

1983/100 ms, resolution = 0.67x0.67x6.48 mm^3). In the later part of the cohort a different anatomical scan was acquired, namely a respiratory-triggered scan with Multi-Vane motion compensation (TR/TE = 2039/100 ms, resolution = 0.62x0.62x3 mm^3, SENSE = 2). Respiratory-triggering was performed using a navigator placed on the diaphragm/liver interface. A coronal IVIM-MRI was acquired with 13 b-values ranging from 0 to 800 $s\,mm^{-2}$ (b-values= 0, 10, 20, 30, 40, 50, 75, 100, 150, 200, 350, 600, 800 $s\,mm^{-2}$, resolution = 3.25x3.25x4.0 mm^3, bandwidth per pixel = 22.9 Hz, TR = 4087 ms, free breathing). Number of squared averages (NSA) for all b-values was equal to 2, except for b=0, 600 and 800 with NSA = 4 and b=200 and 350 with NSA = 3, which resulted in a total scanning time of 6:36 minutes for the IVIM scan-series. Correction for geometric distortions in the IVIM images was performed using a B_0field map (resolution 4.06x4.06x4 mm^3, TR = 630.4 ms, TE_1/TE_2 = 4.6/9.2 ms, flipangle = 30°) [124].

Image analysis

Three-dimensional B-Spline registration on the IVIM series was used to mitigate the respiratory motion within the acquisition (registration between even and odd slices) using an approach introduced by Guyader et al. [142]. Delineation of the tumor was performed on the initial b800 $s\,mm^{-2}$ scan using an automated contouring algorithm based on clustering (ITK-SNAP [98]) and verified by 2 readers (A.S.B. and S.E.H.) using the anatomical T_2W scan (Figure 5.1). For each of the subsequent time points a rigid registration was performed, based on the transverse T_2W, with respect to the baseline scan. Using this registration the initial delineation was propagated to the subsequent time points and manually adjusted if necessary (e.g. for shrinkage of the esophageal circumference). Within this delineation analysis of IVIM-parameters was performed.

The current study investigated different fitting approaches of the bi-exponential (or Bihan) model, as presented in Equation 5.1:

$$S_b = S_0 \cdot \left[(1 - F_p) \cdot e^{-b \cdot D} + F_p \cdot e^{-b \cdot D^*}\right] \tag{5.1}$$

with S_0 representing the signal at b = 0 $s\,mm^{-2}$, F_p the perfusion-fraction, D the diffusion and D^* the pseudo-diffusion. Since conventional DWI analysis is based on mono-exponential fitting of two or three b-values (Equation 5.2), this analysis was also included for reference purposes:

$$S_b = S_0 \cdot e^{-b \cdot ADC} \tag{5.2}$$

with ADC representing the apparent diffusion coefficient. A schematic overview of the different fitting approaches is presented in Figure 5.2.

Table 5.1: Patient characteristics

Characteristic	n	
Gender		
Male	22	(92%)
Female	2	(8%)
Age (at start RT), years (SD)	67.3	(6.7)
Clinical T-stage		
cT1	1	(4%)
cT2	1	(4%)
cT3	22	(92%)
Clinical N-stage		
N0	8	(33%)
N1	12	(50%)
N2	3	(13%)
N3	1	(4%)
Histopathology		
Squamous cell carcinoma	6	(25%)
Adenocarcinoma	17	(71%)
Undifferentiated carcinoma	1	(4%)
Location		
Proximal third of esophagus	0	(0%)
Middle third of esophagus	1	(4%)
Distal third of esophagus	21	(88%)
Gastro-esophageal junction	2	(8%)
GTV (mL)		
Mean (SD)	48.5	(27.5)
Range		6.7 - 126.3

GTV = gross tumor volume.
NOTE - Data presented as counts with percentages in the parentheses unless stated otherwise

The mono-exponential fits were calculated using an in-house written tool using b-values of 0, 200 and 800 $\mathrm{s\,mm^{-2}}$ (approach 1, ADC0) and b-values of 50, 200 and 800 $\mathrm{s\,mm^{-2}}$ (approach 2, ADC50). Since the influence of the pseudo-diffusion (D^*) can typically be neglected for higher b-values, the bi-exponential model using a non-linear least squares fit was fitted in a two-step manner, where the initial fit included a weighted fit of the diffusion (D) and perfusion-fraction (F_p) using all b-values $\geq 200\ \mathrm{s\,mm^{-2}}$ [143,144]. The weight was defined as: $\mathrm{w_b}=\mathrm{signal_b}\cdot\sqrt{\mathrm{NSA}}$, as this accounts for the assumed differences in SNR between the b-values. Subsequently, using all 13 b-values, D^* was fitted using the input-parameters as calculated in the initial fit (approach 3, Bihan-fixed). In addition, the second fit was also performed fitting all parameters and using the initial fitted F_p and D as starting values (approach 4, Bihan-free). In approaches 3 and 4 the following limits were used for D, F_p and

D^* respectively: minima of $1{\cdot}10^{-5}$ $mm^2\,s^{-1}$, 0.0005, $1{\cdot}10^{-3}$ $mm^2\,s^{-1}$ and maxima of $6{\cdot}10^{-3}$ $mm^2\,s^{-1}$, 0.9995, 0.125 $mm^2\,s^{-1}$. In addition, a Bayesian-probability approach was fitted including all 13 b-values, with the probability for the diffusion-weighted signal defined as:

$$P(S|DF_pD^*) \propto \left[\tfrac{1}{2}\sum_b \left(S(b) - S_0 \cdot [(1-F_p)\cdot e^{-b\cdot D} + F_p \cdot e^{-b\cdot D^*}]\right)^2\right]^{\frac{N}{2}} \qquad (5.3)$$

in which N represents the total number of b-values [40]. For all three parameters a prior-distribution was generated including knowledge of expected parameters values for esophageal cancer from previous reports [145, 146] (see inlay-figures in Figure 5.2). An initial 200 'burn-in' samples were used and a total of 3000 generated samples, similarly as described by Barbieri et al [40]. Finally, the output consisted of 3000 combinations of parameter values. Since these parameters are correlated the final values for D, F_p and D^* were estimated by assessing the maximum value in a 3D-histogram of the 'trace', preserving the correlations in the 3000 samples (approach 5, Bayesian).

Statistical analysis

Model performance of each fitting procedure was assessed by the calculation of the coefficient of determination, i.e. R^2. The R^2 value describes to what extend the model fit describes the IVIM datapoints. For $R^2 = 1$ a perfect fit through all data points is found, whereas for a R^2 value below 1, there is a remaining undescribed variance in the data. In this study the same model (Equation 5.1) is used for the IVIM data analysis, for which three different fitting approaches are compared.

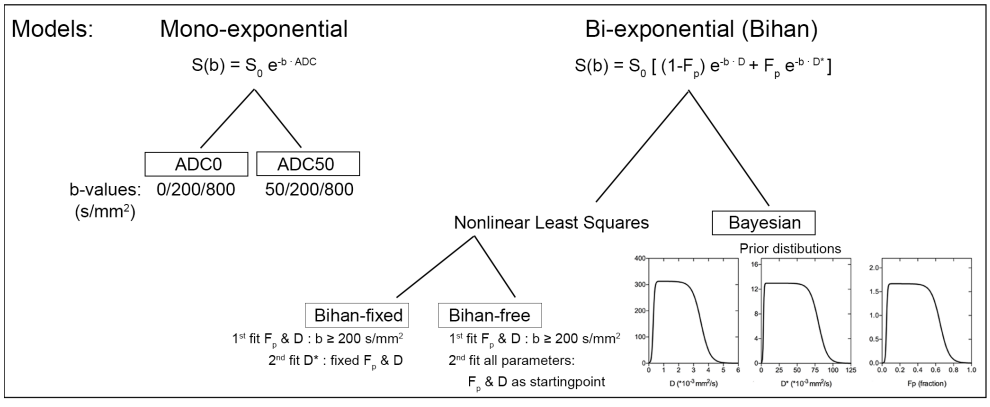


Figure 5.2: A flowchart of the different models and fitting approaches on to fit IVIM-data. The fitting approaches for the Bihan model were calculated in Matlab. Names of all approaches, as referred to in the text, are out-lined. For the Bayesian approach the used input prior distributions are visualized underneath.

Table 5.2: Model comparison (n=47.294 voxels) of the IVIM-based approaches. For reference the values from the conventional ADC models (using 3 b-values) are also included. For all values data is presented as median (25^{th}- 75^{th} percentile).

	IVIM-analysis			Conventional ADC	
	Bihan-fixed	Bihan-free	Bayesian	ADC0	ADC50
Model performance					
R^2	0.96 (0.92 - 0.98)	0.97 (0.94 - 0.98)	0.97 (0.94 - 0.98)	*N.A.*	*N.A.*
Fitted parameters					
D ($\cdot 10^{-3}$ mm^2 s^{-1})	1.77 (1.44 - 2.12)	1.78 (1.33 - 2.22)	1.85 (1.47 - 2.26)	2.58 (2.11 - 3.15)	2.07 (1.75 - 2.46)
D^* ($\cdot 10^{-3}$ mm^2 s^{-1})	22.0 (13.8 - 39.9)	28.9 (11.9 - 73.6)	32.7 (14.4 - 58.5)	*N.A.*	*N.A.*
F_p	0.32 (0.18 - 0.47)	0.33 (0.20 - 0.51)	0.31 (0.20 - 0.45)	*N.A.*	*N.A.*

Patient and treatment-related characteristics were described as count with percentages or mean with standard deviation (SD) as appropriate. All calculations were performed in Matlab (The Mathworks Inc, Natick, MA) and all image registrations in Elastix (Version 4.7) [147], if not stated otherwise. Bland-Altman density plots were used to illustrate the differences between the various IVIM fitting approaches, including lines at the mean and upper and lower limits of agreement (LoA) ($\pm 1.96 \times$ SD).

Results

Patient and tumor characteristics

Patients had a mean age ($\pm$SD) at the start of radiotherapy of 67.3$\pm$6.7 years, and most of them presented with a clinical T3-stage tumor (22/24). Further patient characteristics are presented in Table 5.1. The 140 IVIM-series originating from 24 patients resulted in a total of 47.294 suspected tumor voxels to be analyzed.

Fitting results

Table 5.2 summarizes the results of the four parameters (D, F_p, D^* and R^2) over all separate voxels within the delineation of the tumor for each time point and each patient, presenting the data as median (25^{th} - 75^{th} percentile). To visualize the differences in model fits and parameter maps, Figures 5.3-5.4 show the model fits through two tumor-voxels and the parameter maps D, F_p and D^* for one tumor. The various model fits using all 13 b-values can appear to be very similar (Figure 5.3a), but can also differ (Figure 5.3b) within two different voxels from the same tumor. Similar to Figure 5.3, a comparison of the different approaches within a tumor (Figure 5.4) shows an overall agreement within the fitted parameters, with similarities in heterogeneity within the tumor. However it is also apparent that D^* shows larger discrepancies between the three approaches compared to D and F_p.

Computational time

Calculation of the mono-exponential models was performed on the entire image space, while for the IVIM-based models calculations were performed on the tumor delineation with a 15 mm surrounding margin. The mono-exponential models had a typical computation time of $<$5 seconds per patient per time point, while for $\approx$3000 voxels (delineation of 11 mL with additional margin) calculation of all three IVIM-fitting approaches was comparable, taking approximately 15-20 minutes per approach (parallel calculations on typically 4 cores).

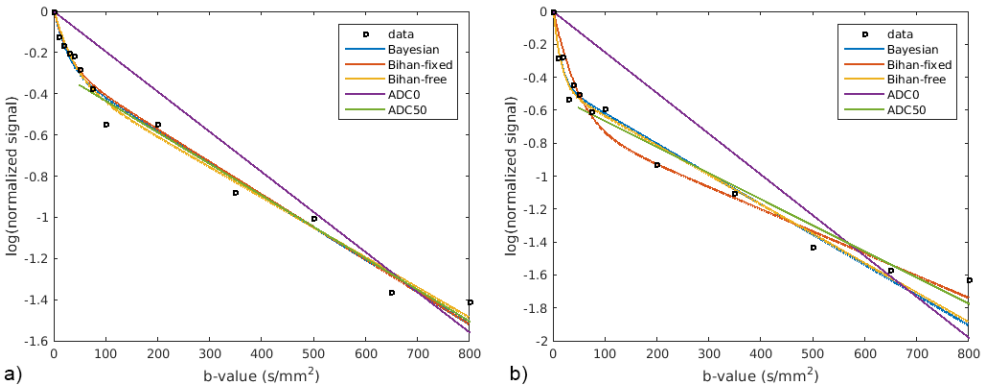


Figure 5.3: a) Signal decay of one voxel as function of b-value. Different lines represent different fitting algorithms. ADC50 is more in line with the bi-exponential models due to the fact that large part of the perfusion-dominated part is excluded in this fit. Furthermore, although most algorithms appear very similar, small visual differences can indicate large differences in D^* (41.8, 38.6 and 28.8 $*10^{-3}$ $mm^2\,s^{-1}$ for the Bayesian, Bihan-fixed and Bihan-free approaches, respectively). **b)** Similar to **(a)** a signal decay of a different voxel in the same tumor.

Voxel-based Bland-Altman analysis

To visualize the differences in the fitting results of the three parameters (D, F_p, D^*) and R^2 between the three fitting approaches on a voxel-based level Bland-Altman density plots of all voxel values were created (Figure 5.5). This data includes voxels of different tumor types, different time points during nCRT, and of both of good and of poor responders. The R^2 density-maps (Figure 5.5j-l) visualize that main part of all fitted voxels result in high model performance ($R^2 > 0.95$ within the majority of voxels). However, it is demonstrated that both the Bihan-free and Bayesian approaches have an offset towards higher R^2 values compared to the Bihan-fixed approach, since both the mean and the majority of points lie below zero (Figure 5.5j-k). When comparing the Bayesian and Bihan-free approaches, these approaches seem to behave very similar, although there is a small offset away from the mean line (Figure 5.5l).

Parameter D (Figure 5.5a-c) has a relatively dense hotspot of values located around $2{\cdot}10^{-3}$ $mm^2\,s^{-1}$, indicating that all approaches find D values in similar range. Furthermore, the anisotropic shape indicates that the 'per-voxel-agreement' does not depend highly on the fitted value for D. The Bihan-fixed and Bayesian approaches behave the most similar, indicated by the smallest LoA. Generally the Bayesian approach resulted in smaller D values compared to Bihan-free (distribution skewed towards negative values). However, the median (Table 5.2) and mean values were higher for the Bayesian approach. The light-blue area crossing the upper LoA in both Figure 5.5a and c indicates that the Bihan-free approach converges to very small values in some voxels, in contrast to both the Bihan-fixed and Bayesian approaches.

For F_p (Figure 5.5d-f) most voxels resulted in values between 0.1 and 0.5 and that the agreement between approaches decreases for larger F_p. Comparing the Bihan-fixed approach with both the Bihan-free and Bayesian approaches, both the Bihan-free and Bayesian approaches showed smaller values for F_p since large part of the voxels lie above zero. The small hotspot for low F_p arising in Figure 5.5e, just below zero, represents voxel values that were found to be close to zero in Bihan-fixed while the Bayesian approach fitted higher (more realistic) values for F_p. The light-blue region in Figure 5.5d, underneath the lower LoA, indicates a model instability of the Bihan-free approach, since these voxels have fitted values close to the upper fitting boundary compared to regular fitted values in the Bihan-fixed approach.

All figures for D^* (Figure 5.5g-i) show wide LoA when compared to the range of the majority of points on the x-axis. This visualizes the large variety found in fitted D^*

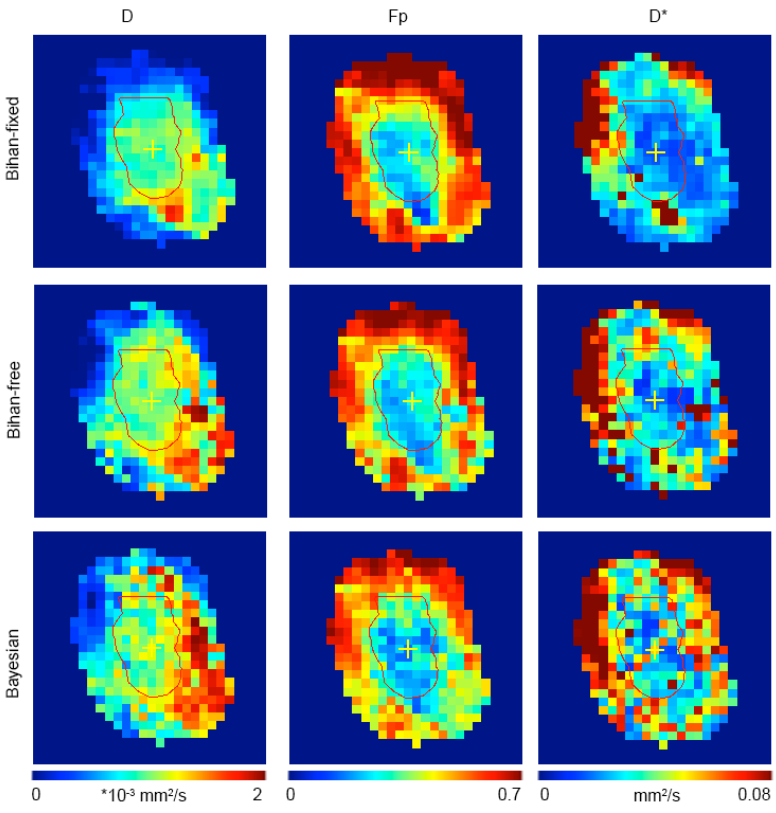


Figure 5.4: For the same location in the tumor as in Figure 5.1, the different parameter maps D, F_p and D^* are shown (sagittal plane) for both the Bihan-fixed, Bihan-free and Bayesian approaches. The colorbar indicates the scaling of the different parameters. The tumor delineation is depicted in red, with a linked voxel indicated by the yellow crosshair.

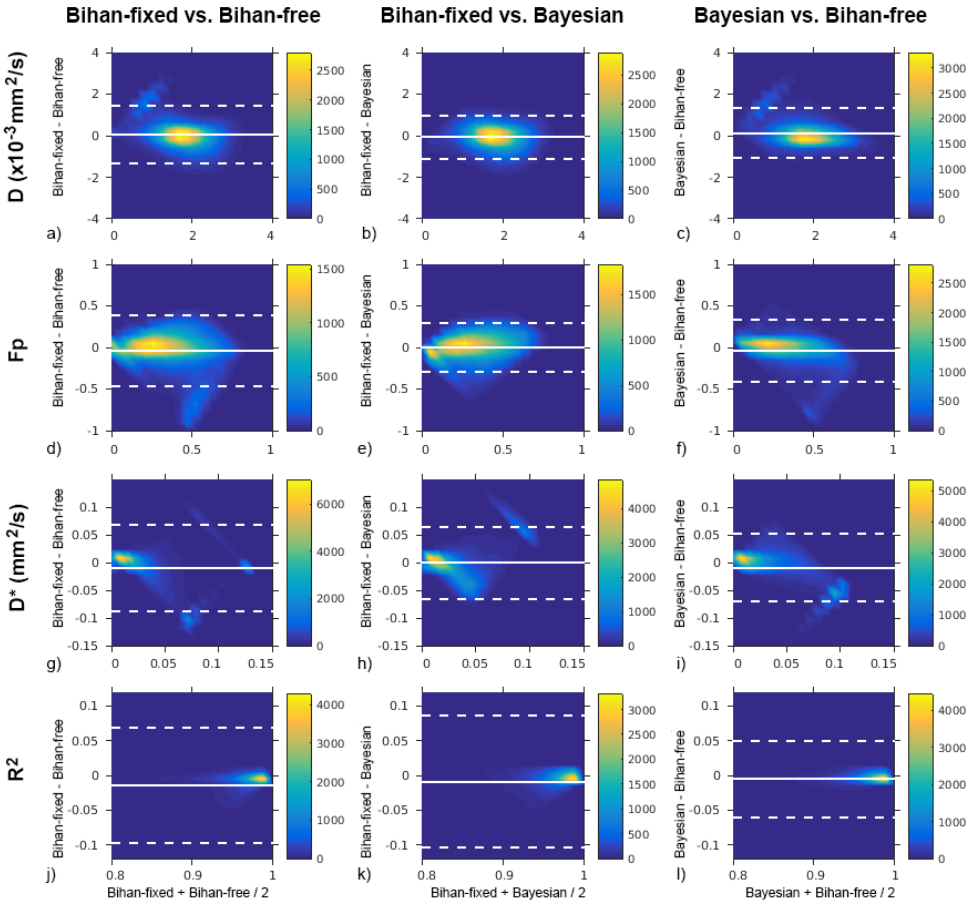


Figure 5.5: Bland-Altman density-maps of each fitted parameter (D, F_p and D^*) and R^2 value for all voxels (N= 47.294). The first column represents a comparison between the Bihan-fixed and the Bihan-free approaches, the second column between the Bihan-fixed and the Bayesian approaches, and the third column between the Bayesian and the Bihan-free approaches. For each Bland-Altman plot the mean and $\pm 1.96 \times$ SD lines are shown (solid and dashed respectively).

values in all approaches compared to D and F_p. It furthermore indicates that D^* is the most sensitive for the chosen fitting approach. In Figure 5.5h, the light-blue area towards 0.05 $\mathrm{mm^2\,s^{-1}}$ (x-axis) indicates voxels with a very poor per-voxel-agreement since the difference between the models equals the mean of both models. Furthermore, comparing the Bihan-free approach with both the Bihan-fixed and Bayesian approaches it is apparent that the core is tilted above zero, indicating that Bihan-free often resulted in lower D^* values.

Global comparison over the time points

To summarize the voxel-based comparison for the different model fits all three parameters (D, F_p, D^*) were presented as boxplots, where the voxels are separated for the different time points (Figure 5.6). As expected both mono-exponential models resulted in highest diffusion values, although using 50 $\mathrm{s\,mm^{-2}}$ as initial b-value (ADC50) approaches similar values compared to the bi-exponential models. For reference purposes the literature values from recent publications concerning esophageal cancer are also included for each parameter [145, 146]. Both D and F_p show clear increasing trends over time, which is in accordance with the presented literature values. The larger whiskers (10-90$^{\mathrm{th}}$ percentile) found for the Bihan-free approach in all parameters, underlines the larger variety in fitted values for this approach as was also seen in Figure 5.5. These error bars end at the maximum value allowed for this type of fit.

Discussion

In this study, 47.294 voxels from IVIM-data series of 24 patients with esophageal cancer were analyzed to assess the influence of using three different fitting approaches of a bi-exponential model. Although a high model performance was found for all approaches, influence of the different approaches on the fitted parameters was observed showing different behavior in fitted values. It is expected that these differences influence the accuracy and precision in longitudinal studies for response assessment in which IVIM will be used. Therefore, we argue for the explicit reporting of specifications of fitting approaches used in IVIM studies to be able to compare outcomes between studies.

The high R^2 values found in the different approaches imply that all fits represent the IVIM-datapoints well. However, although high R^2 values were found for most voxels, large variations can occur for the three fitted parameters comparing the different approaches. In order to accurately assess response based on differences in IVIM-data between time points [66, 123], it is important to have a well-described fitting method. Using simulated data, Barbieri et al. [40] showed similar trends as the current study between the Bihan-fixed and Bayesian approaches, namely, that the Bihan-fixed approach resulted in lower D^* with a smaller range and a tendency towards higher F_p values (at lower signal-to-noise ratios). The agreement between the fitted parameters of their simulated data and our data could hint towards the Bayesian approach as the preferred fitting approach since this approach showed superior performance in the simulated data. Furthermore, our data showed a minimal offset towards higher R^2 values for the Bayesian approach compared to the Bihan-fixed approach.

The larger variety for the different parameters apparent in Figure 5.6, comparing the

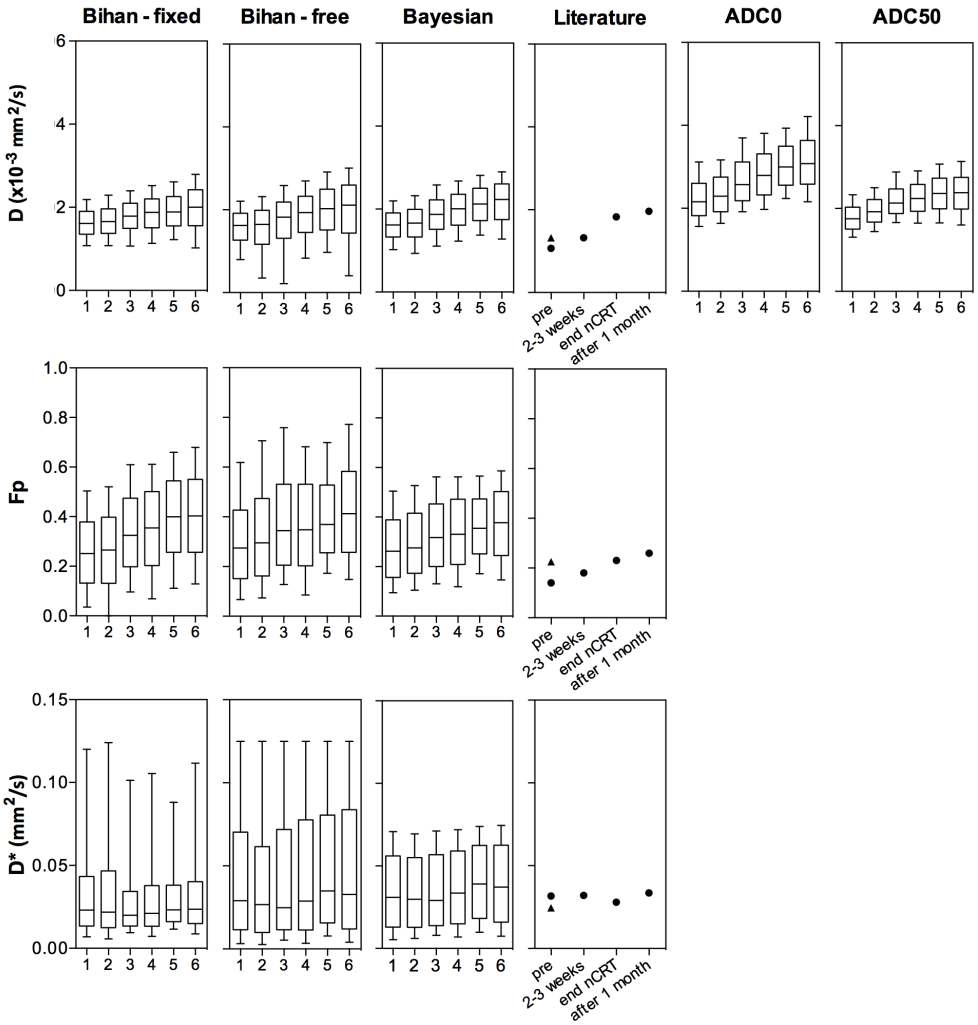


Figure 5.6: Three IVIM parameters D, F_p and D^* (row 1-3 respectively) plotted for the different fitting approaches, in which all voxels (N= 47.294) are separated for de six time points (x-axis). For the diffusion D also the reference models ADC0 and ADC50 are shown. All the boxplot represent the interquartile range including the median line, the whiskers represent the 10-90$^{\text{th}}$ percentile values. For each parameter also literature values are depicted from two publication concerning IVIM-fits on esophageal cancer [145, 146], with the timing of the scans with respect to nCRT depicted on the x-axis.

Bihan-free to the Bihan-fixed approach, can be explained theoretically. Model parameters F_p and D were fitted on the higher diffusion weights (b $\geq$ 200 s mm^{-2}) and set fixed in the final fit for D^* in the Bihan-fixed approach, while in the Bihan-free approach these values were only used as starting values. This property provides more freedom in fitting, resulting in a larger variety of fitted values (see also Figure 5.3b).

Furthermore, since the acquired diffusion-weights were not linearly spread but included a majority of data points in the lower segment of diffusion-weights (to be able to fit the pseudo-diffusion), the weight of the R^2 calculation tends towards a good fit through those points. Therefore, the Bihan-fixed approach resulted in slightly lower R^2 values, as seen in Figure 5.5, since this approach is not allowed to optimize freely in the lower segment of b-values.

Focusing on the range of the IVIM parameters found in Figures 5.5 and 5.6 it is important to keep in mind that this range is also affected by tumor- and treatment-related differences (e.g. treatment induced changes, tumor type, and good or poor response to treatment), as no discrimination was made based on tumor type or response (Figure 5.6) and time points (Figure 5.5). Therefore a smaller range in fitted values in a certain approach does not necessarily imply superiority over other approaches. The slightly lower estimated F_p values by the Bayesian-probability fit were more in line with literature values reported for esophageal cancer (Figure 5.6) [145,146] and various other tumor sites [40,148,149], compared to both the Bihan-fixed and the Bihan-free approaches. One of the explanations is that both the fixed and the free algorithms are (initially) driven by b-values $\geq 200\ \mathrm{s\,mm^{-2}}$ in which especially the highest b-value measurements lie close to the noise floor. This could artificially increase the absolute signal intensity, as the Rician noise regime applies to this situation [150]. In this situation an underestimation in D and overestimation in F_p can be expected. Furthermore, as Figure 5.3 visualizes, even a small difference in fit can result in a large difference in especially D^*, underlining the importance of a robust method with high precision and accuracy. Current data showed, as expected, that a mono-exponential fit using $50\ \mathrm{s\,mm^{-2}}$ as initial b-value instead of using $0\ \mathrm{s\,mm^{-2}}$ revealed more resemblance to diffusion values as estimated by the bi-exponential model. The potential added value of fitting IVIM-data providing both perfusion and diffusion information in relation to response prediction, compared to solely using mono-exponential fitting, remains a topic of future research.

Recently, multiple studies were published using IVIM-MRI for esophageal cancer in relation to early diagnosis, tumor staging and response prediction [128, 145, 146]. A preliminary study comparing IVIM-MRI of the healthy esophagus versus esophageal cancer reported F_p values of 0.64 and 0.48, respectively [128]. These values were higher than the values found in our study. However in the preliminary study no specifications were reported of the bi-exponential fitting approach. Two studies publishing IVIM-analysis (as included in Figure 5.6) reported values estimated in only 1 or 3 tumor slices [145, 146]. In one of these studies [146] no specifications were given on the fitting approach. These studies reported lower values of perfusion fraction F_p and D compared to the approaches in our study. Similar to the current study, the presented values from Zheng et al. [145] included both good and poor responders and the first 3 measurements cover a similar time range during the course of treatment. In this time

range, a similar increasing time trend of both D and F_p during the course of treatment was reported. It is promising that the observed increasing time trend in the perfusion parameter F_p is in accordance to the perfusion as measured by DCE-MRI [123], in which in general a mild to strong increase (depending on response) was observed in the AUC over the course of treatment. As also stated in earlier studies [40, 151], the current study confirms that the different parameters are dependent on the chosen fitting approach, as can be appreciated from Figures 5.5 and 5.6. Comparison with literature values is therefore challenging since specifications on used fitting approaches are often inadequate or sometimes even absent.

A few limitations apply to this study. First, it could be further investigated how large the influence is of the choice of b-values [152]. A large number of low b-values were included (8 values $\leq 100\ \mathrm{s\,mm^{-2}}$) in the imaging protocol of the current study to enable accurate fitting of the pseudo-diffusion and less scans were acquired in the higher segment (3 values $> 200\ \mathrm{s\,mm^{-2}}$). Second, the diffusion-weighted scans were acquired during free breathing. Subsequently, registration of motion was performed to compensate for motion within the scans [142]. The impact of using respiratory-triggered or breathhold acquired diffusion-weighted imaging instead of free breathing acquisition on calculated model parameters is under debate [153, 154], as the presence of motion within the scans could artificially increase the estimated F_p. Lastly, since this study was a methodological study focusing on differences in behavior of voxel-level, the overall spread in the calculated parameters is affected by treatment induced changes, tumor type and possible response to treatment. Since no ground-truth is available in clinical data, and to our knowledge no IVIM-phantom exists, this study is limited to a comparison between the different fitting approaches. In future studies, repeated measurements without intermediate treatment are of interest to judge which of the fitting methods gives the most reproducible result. Alternatively, the performance of the IVIM parameters to assess treatment response can be compared between the fitting methods. In this approach the model that delivers parameters of the highest discriminative power should be selected for inclusion in a (multi-parametric) treatment assessment protocol.

In conclusion, although no large discrepancies were found in model performance, it was found that large differences in behavior and variance occur between different fitting approaches comparing the fitted diffusion and perfusion parameters. The Bihan-free approach showed to have numerous outliers in the D^* fit in which case the fit clipped at the maximum value, based on which we least advise to use this method. Both the Bayesian and Bihan-fixed approaches resulted in high R^2 values and did not show the clipped behavior as observed in the Bihan-free approach. The time trends found in D and F_p for these approaches are promising in light of the potential to predict response in future research. The differences in the observed fitting parameters of the different fitting approaches underline the importance to use uniform fitting procedures and

data analysis in multi-center trials. Furthermore, it is essential to include explicit reporting of fitting specifications when publishing results on IVIM-data to be able to compare results between studies.

Chapter 6

Quantification of variations in intra-fraction motion of esophageal tumors over the course of neoadjuvant chemoradiotherapy based on cine-MRI

The following chapter is based on:

Sophie E. Heethuis, A.S. Borggreve, Lucas Goense, Peter S.N. van Rossum, S. Mook, Richard van Hillegersberg, Jelle P.Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier

Physics in Medicine and Biology (2018) Vol. 63 (14):145019

Abstract

Purpose
To noninvasively quantify variation in intra-fraction motion of esophageal tumors over the course of neoadjuvant chemoradiotherapy (nCRT) using 2D cine-magnetic resonance imaging (MRI) series.

Material and Methods
Patients treated with nCRT for esophageal cancer underwent 6 MRI scans. Scans were acquired prior to the start of nCRT, followed by weekly MRI scans during nCRT. Cine-MRI series were acquired in the coronal and sagittal plane (≈1.6 Hz). To be able to quantify intra-fraction motion over a longer time period, a second cine-MRI series was performed after 10 minutes. Tumor motion was assessed in cranio-caudal (CC), anterior-posterior (AP) and left-right (LR) direction. Motion patterns were analyzed for the presence of deep inhales and tumor drift.

Results
A total of 232 cine-MRI series of 20 patients were analyzed. The largest tumor motion was found in CC direction, with a mean peak-to-peak motion of 12.7 mm (standard deviation [SD] 5.6), followed by a mean peak-to peak motion in AP direction of 3.8 mm (SD 2.0) and in LR direction of 2.7 mm (SD 1.3). The CC intra-fraction tumor motion can differ extensively between and within patients. Deep inhales were present in 6 of 232 scans (3%). After exclusion of these scans, mean CC peak-to-peak motion was 12.3 mm (SD 5.2). Correction for tumor drift showed a further reduction to 11.0 mm (SD 4.6). Despite correction for tumor drift, large variation in tumor motion occurred within patients during treatment. Mean tumor drift during the 10 minute interval between the 2 series was 1.5 mm (SD 1.8), with a maximum of 11.6 mm.

Conclusions
Intra-fraction tumor motion was found to be highly variable between and within patients with esophageal cancer over the course of nCRT. Correction for deep inhales and tumor drift reduced peak-to-peak motion. The stochastic nature of both deep inhales and tumor drift indicates that real-time tumor motion management during radiotherapy is a prerequisite to safely reduce treatment margins.

Introduction

Neoadjuvant chemoradiotherapy (nCRT) followed by surgical resection is widely accepted as standard treatment with curative intent for patients with resectable esophageal cancer [11,58,61]. Due to the presence of positioning uncertainties and respiratory motion, substantial CTV-to-PTV margins (typically ranging from 10 mm to 15 mm) are currently applied [11]. Smaller CTV-to-PTV margins would result in a reduced radiation dose to the surrounding healthy tissues, and would thereby reduce short- and long-term complications. Furthermore, smaller margins could be used to increase the dose to the tumor, while maintaining the dose constraints for healthy tissues that are currently used. In order to develop highly conformal treatment plans with adequate tumor coverage, the irradiation field needs to be adapted to tumor movement. As such, information about (variations in) tumor movement over the course of nCRT is desired.

To date, several imaging modalities have been used to assess tumor motion in esophageal cancer patients. Drawbacks of these studies using four dimensional computed tomography (4DCT) to quantify esophageal tumor motion [47,155–157] are the additional exposure to radiation and the need for invasive placement of fiducial markers. Furthermore, 4DCT studies only provide information on the average tumor motion as observed during CT acquisition, but not on variations in the tumor motion during scanning. This is due to the reconstruction of 4DCT scans in which the tumor motion is binned in a defined number of respiratory phases or amplitudes after acquisition, resulting in the reconstruction of only one breathing cycle in 4DCT. Therefore it is not possible to measure tumor drift with 4DCT, as measurement of this process requires reconstruction of multiple breathing cycles.

Esophageal tumor motion can also be visualized with magnetic resonance imaging (MRI). In cine-MRI, subsequent 2D images are acquired with a high temporal resolution, which is widely used to study motion patterns in different tumor sites [158–160]. A mean cranio-caudal peak-to-peak motion of 13.3 mm with a wide range of 2.7 to 24.5 mm was found using cine-MRI, in which the mobility of distal tumors was generally larger than that of more proximal tumors [48].

In previous studies either four-dimensional cone-beam CT (4D-CBCT) was used over the course of treatment, to quantify variations in tumor motion, or MRI was used at a single time point prior to treatment. As mentioned above, the disadvantage of 4DCT imaging is that it does not give information about possible tumor drift. Therefore, the purpose of our observational study was to quantify variations in intra-fraction esophageal tumor motion over the course of nCRT within the same patient, using weekly cine-MRI series. Insight in motion patterns of esophageal tumors could be used to determine which motion management strategy (e.g. tumor tracking, tumor gating, dynamic planning) [55, 57] may be the most appropriate for on-line MRI-guided

radiotherapy [54, 56]. Furthermore, if tumor motion prior to treatment is a reliable indicator for tumor motion during treatment, personalized PTV margins could be used in future clinical practice.

Material and methods

Study population

Patients with histologically confirmed esophageal cancer who were planned to receive nCRT followed by surgery were eligible for inclusion. Neoadjuvant chemoradiotherapy consisted of 5 weeks of treatment, involving weekly intravenous administration of carboplatin and paclitaxel with concurrent radiotherapy (41.4 Gy in 23 fractions of 1.8 Gy) according to the CROSS regimen [11]. Standard diagnostic work-up consisted of endoscopy with biopsies of the primary tumor site, as well as integrated ^{18}F-FDG-PET/CT for clinical staging. Patients with contraindications for MRI were not eligible for inclusion. The study was approved by the local medical ethics committee and written informed consent was obtained from all patients.

MRI acquisition

MRI images were acquired at 6 time points; prior to the start of nCRT, followed by weekly MR scans during the course of nCRT. All images were acquired on a 1.5 T Philips Ingenia (Best, the Netherlands), using the Anterior/Posterior (28 channel) receive coils. Patients were positioned in supine position with both arms next to the body. Respiratory-triggered sagittal and transverse anatomical T_2-weighted scans were acquired with a multi-slice turbo spin echo sequence (tT_2W: TR/TE = 1983/100 ms, sT_2W: TR/TE = 1432/100 ms, resolution = 0.67x0.67x4 mm^3). These scans served as reference to plan 2 cine-MRI series: one coronal and one sagittal cine-MRI

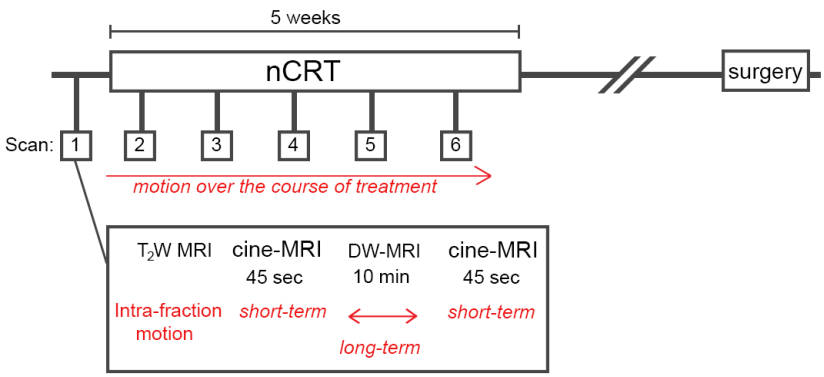


Figure 6.1: Overview of treatment and scanning timeline including a schematic representation of terminology used for indication of intra-fraction motions quantified in this study.

series To prevent substantial out-of-plane motion the scanning plane of the coronal cine-MRI series was angulated, such that it was aligned with the largest extent of the tumor in the sagittal T_2-weighted scan. For the sagittal cine-MRI series, the scanning plane was positioned through the center of the tumor, which was identified using the transverse T_2-weighted scan. The images were both acquired with a resolution of 2.01x2.01 mm and 5 mm slice thickness, during 46 seconds of free breathing (1.6 Hz) for the coronal and 43 seconds (1.7 Hz) for the sagittal series. In order to quantify the difference in intra-fraction tumor motion and drift within one scanning session, both coronal and sagittal cine-MRI series were acquired twice during each scanning session: the first series was initiated after approximately 20 minutes of scanning and the second series after scanning a diffusion-weighted (DW) MR series (10 minutes). These DW-MR scans included a b-value of 800 $s\,mm^{-2}$ (SPIR fat suppression, resolution = 3.25x3.25x4 mm^3, bandwidth per pixel = 22.9 Hz).

Motion characterization

The primary tumor was delineated on the initial b800 $s\,mm^{-2}$ DW-MRI, using an automated contouring algorithm based on clustering (ITK-SNAP [98]). Afterwards, this delineation was verified using the transverse T_2W scan (tT_2W). For each of the subsequent time points (2 to 6) a rigid registration between the tT_2W scan and the tT_2W baseline scan was performed in Elastix (Version 4.7) [147]. Using this registration, the tumor delineation of the first time point was propagated to the other time points (2 to 6) and manually adjusted if necessary (e.g. in case of shrinkage of the esophageal circumference). Next, all tumor delineations were propagated to the corresponding 2D cine-MRI series. An optical flow algorithm (RealTITracker [161, 162]) was used to quantify tumor motion on the cine-MRI series. Corresponding motion vector fields were calculated for each frame with respect to a reference frame and evaluated within the delineated tumor volume. The reference frame was manually selected from 6 consecutive frames in such a way that the shape of the tumor resembled the tumor delineation as performed on the initial b800 scan as closely as possible. Motion was evaluated in cranio-caudal (CC), left-right (LR) and anterior-posterior (AP) directions (Matlab, The Mathworks Inc, Natick, MA, USA). The angulation of the coronal scanning plane was taken into account for the calculation of the CC-motion, such that the true CC-motion was calculated. Peak-to-peak amplitudes of tumor motion were calculated in all three directions. Furthermore, the smallest range to include 95% of the data points (excluding 5% most extreme data points) was calculated (C95). This C95 is a more robust measure for tumor motion as it excludes outliers caused by sporadic movement (e.g. sneezing, hiccups) as seen on the cine-MRI series. CC-motion patterns analyses were divided into short- and long-term intra-fraction motion (see Figure 6.1 for schematic overview of timeline and terminology).

Short-term intra-fraction motion was defined as the motion occurring within one cine-MRI series of 45 seconds, in which the presence of deep inhales and tumor drifting or

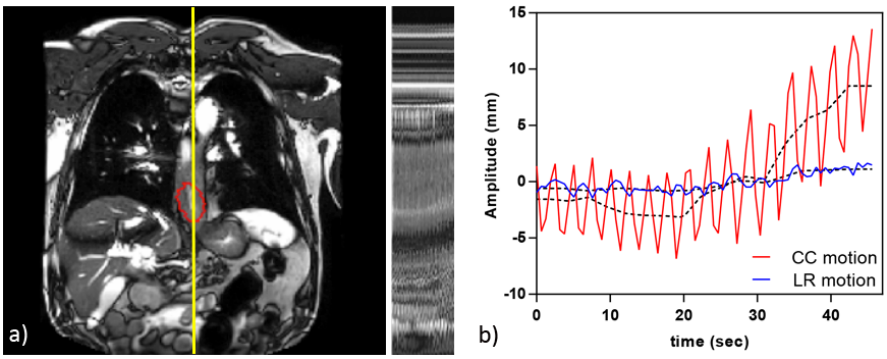


Figure 6.2: a) A cine-MRI scan of a patient with the tumor delineation in red. On the right side the 75 time frames are visible corresponding to the location of the yellow line in the image. **b)** Amplitude of the tumor motion in cranio-caudal (CC, red) and left-right direction (LR, blue). The black dashed lines show the estimated drift lines.

irregular motion patterns were quantified. To estimate the effect of deep inhales on the motion curves, the minimum value (0^{th} percentile, P0) and 10^{th} percentile (P10) were compared. No large P10-P0 differences were expected in a regular breathing curve. However, in a curve in which a deep inhale occurs (presenting as a negative peak in a motion curve), a large difference in P10-P0 was expected. To determine the presence of tumor drift in motion curves, drift lines were estimated. These drift lines were defined as the mid-position (mean) between consecutively found local maxima and minima in the motion curve. This drift line was subtracted from the motion curve, after which the difference in peak-to-peak and C95 amplitude for both curves functioned as a measure for the presence of a drift.

Long-term intra-fraction motion was defined as the motion which occurred within the 10 minute time-interval between the 2 cine-MRI series. This was quantified as the shift between the initial drift lines (first 6 data-points) of both the first and second cine-MRI series. This can be interpreted as a conservative surrogate for the displacement of the esophageal tumor between set-up of the patient and final dose delivery. In which a positive long-term drift represents a movement in cranial direction.

Statistical analysis

This study was of descriptive nature and therefore no formal sample size calculation was performed. All visual representations were generated using GraphPad Prism 6.07 software (GraphPad Software, La Jolla California USA) and statistical analysis was performed in SPSS Statistics version 22 (IBM Corp., Armonk, NY, USA). The mean and standard deviation (SD) of the peak-to-peak and C95 amplitude for the whole group were calculated.

Table 6.1: Patient characteristics

Characteristic	n	
Gender		
Male	20	(100%)
Age (at start RT), years (SD)	66.6	(6.9)
Clinical T-stage		
cT2	1	(5%)
cT3	19	(95%)
Clinical N-stage		
N0	7	(35%)
N1	9	(45%)
N2	3	(15%)
N3	1	(5%)
Histopathology		
Squamous cell carcinoma	4	(20%)
Adenocarcinoma	15	(75%)
Undifferentiated carcinoma	1	(5%)
Location		
Proximal third of esophagus	0	(0%)
Middle third of esophagus	2	(10%)
Distal third of esophagus	15	(75%)
Gastro-esophageal junction	3	(15%)
Volume planning GTV (ml)		
Mean (SD)	50.9	(25.5)
Range		19.0 - 126.3

GTV = gross tumor volume.
NOTE - Data presented as counts with percentages in the parentheses unless stated otherwise

Results

Study population

A total of 25 consecutive patients were included, with MR scans acquired between April 2016 and October 2017. Five patients were excluded for analysis due to small tumor volume (<7 mL, n=1), poor tumor visibility on MRI (n=3) or withdrawal of informed consent (n=1), resulting in a total of 20 patients suitable for analysis. In 3 of the included 20 patients, one of the weekly MR-scans during nCRT was incomplete or missing due to the condition of the patient. In 2 patients the second cine-MR series was not available at one time point. This led to a total amount of 232 acquired cine-MRI series for the total study population. All patients were male (20/20) with a mean age of 66.6 years (SD 6.9). Two patients had a mid-esophageal tumor (patient 1 and 2), 15 patients a distal esophageal tumor (patients 3-17), whereas the remaining 3

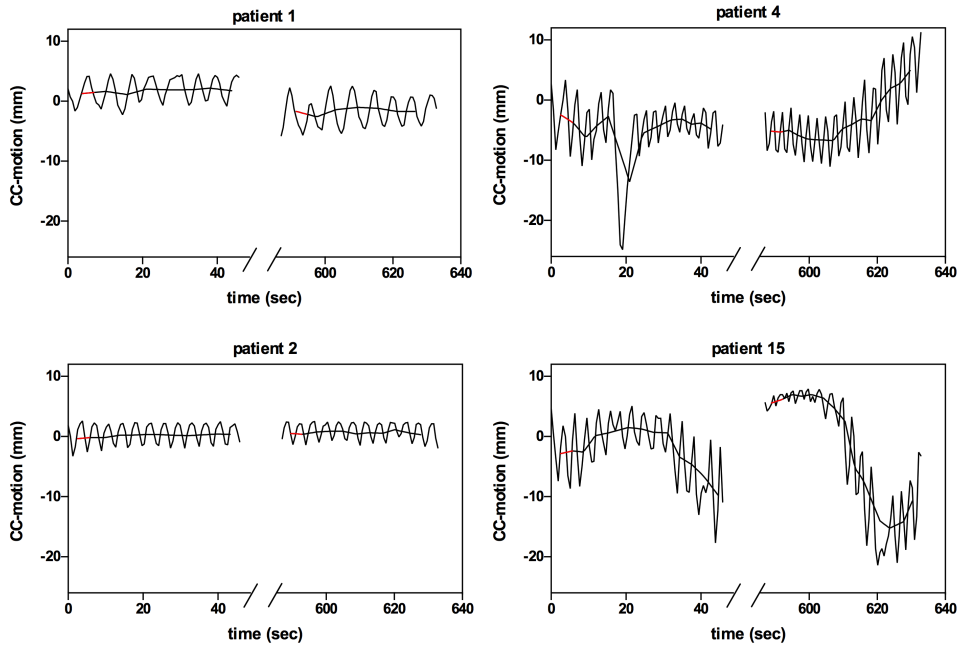


Figure 6.3: Four examples of motion curves with different behavior are visualized. Patient 1: a long-term drift between series. Patient 2: very regular motion. Patient 4: a deep inhale and short-term drift. Patient 15: irregular motion behavior in both series. The first 6 data-points in each drift line, indicated in red, were used to quantify the long-term drift.

patients had a gastroesophageal junction tumor (patients 18-20). Patient and tumor characteristics are summarized in Table 6.1.

Tumor motion

A coronal cine-MRI frame together with the estimated CC and LR motion is shown in Figure 6.2. This figure shows excellent visualization of the thorax and abdomen at 1.5T, enabling motion characterization. The largest intra-fraction esophageal tumor motion was found in CC direction with a mean peak-to-peak motion of 12.7 mm (SD 5.6, range 4.4-33.7). The mean AP peak-to-peak motion was 3.8 mm (SD 2.0, range 0.7-10.2). The least motion was seen in LR direction, with a mean peak-to-peak of 2.7 mm (SD 1.3, range 0.7-8.3). The calculated peak-to-peak and C95 in the CC direction for the whole group is summarized in Table 6.2. Since CC-motion was 22% less on average when calculated on sagittal images, analyses were only performed on the coronal cine-MRI series.

Typical CC tumor motion curves are presented in Figure 6.3. Curves with long-term intra-fraction drift (patient 1) and short-term intra-fraction drift (patient 4), as well

as regular (patient 2) and irregular motion behavior (patient 15) are depicted. The variety of tumor motion for the entire patient group is visualized in Figure 6.4a, in which the peak-to-peak amplitude is shown in each patient for the 2 scans for each time point separately. As illustrated by Figure 6.4a, the average intra-fraction tumor motion can differ extensively between patients, and large variations of intra-fraction motion are seen within some patients. Furthermore, it indicates that amplitude prediction for tumor motion throughout the treatment course based on the initial two baseline series (blue line represent mean initial amplitude), creates a risk of severe overestimation (patient 7) or underestimation (patient 4).

Short-term intra-fraction motion

Analysis of the CC motion patterns based on the P10 and P0 revealed presence of deep inhales in 6 of the 232 series (3%). In these 6 cine-MRI series, the magnitude of the tumor motion during these inhales was 2 times larger than that of the remaining tumor motion, and the P10-P0 value was more than 3 times the standard deviation above the mean of all 232 series. Results of the short-term drift correction, using the drift lines within the series, are visualized in Figure 6.4b. A mean reduction in peak-to-peak amplitude of 1.3 mm (SD 1.8) was found. The largest short-term drift was present in patient 15 (-12.4 mm), which was verified by visual inspection of the cine-MRI series and is also depicted in Figure 6.3 (second series). Since a drift line correction does not compensate for deep inhales, the 3 large amplitudes in the CC motion of patient 7 - caused by 3 deep inhales - remained. As the presence of these deep inhales obscured the effect of the drift line correction, cine-MRI series including deep inhales were excluded. Peak-to-peak and C95 motion amplitudes, as well as the corresponding drift-corrected amplitudes after exclusion of series with deep inhales are presented in Table 6.2. A mean peak-to-peak motion of 11.0 mm (SD 4.6, range 4.1-26.6) was found after exclusion of deep inhales and drift correction. The decrease in peak-to-peak motion after deep inhale exclusion and drift correction is also visualized in Figure 6.5. After drift correction, 49% of the scans showed a maximum peak-to-peak amplitude smaller than 10 mm, compared to 38% before drift correction. Before any correction the peak-to-peak amplitude of 95% of the fractions fell within 24.6 mm, this measure of intra-fraction motion was reduced with 5.2 mm if deep inhale exclusion and drift correction was applied.

Long-term intra-fraction motion

The quantification of shift in CC motion at the start of the drift lines from both the first and second cine-MRI series indicated the presence of drift within the intermediate 10 minutes between the acquisitions of the two series. One time point was excluded due to a table movement between the 2 series. A mean long-term drift (absolute values) of 1.5 mm (SD 1.8) was found, with a maximum long-term drift of 11.6 mm. For each patient the long-term drift is visualized for all 6 time-points in Figure 6.4c.

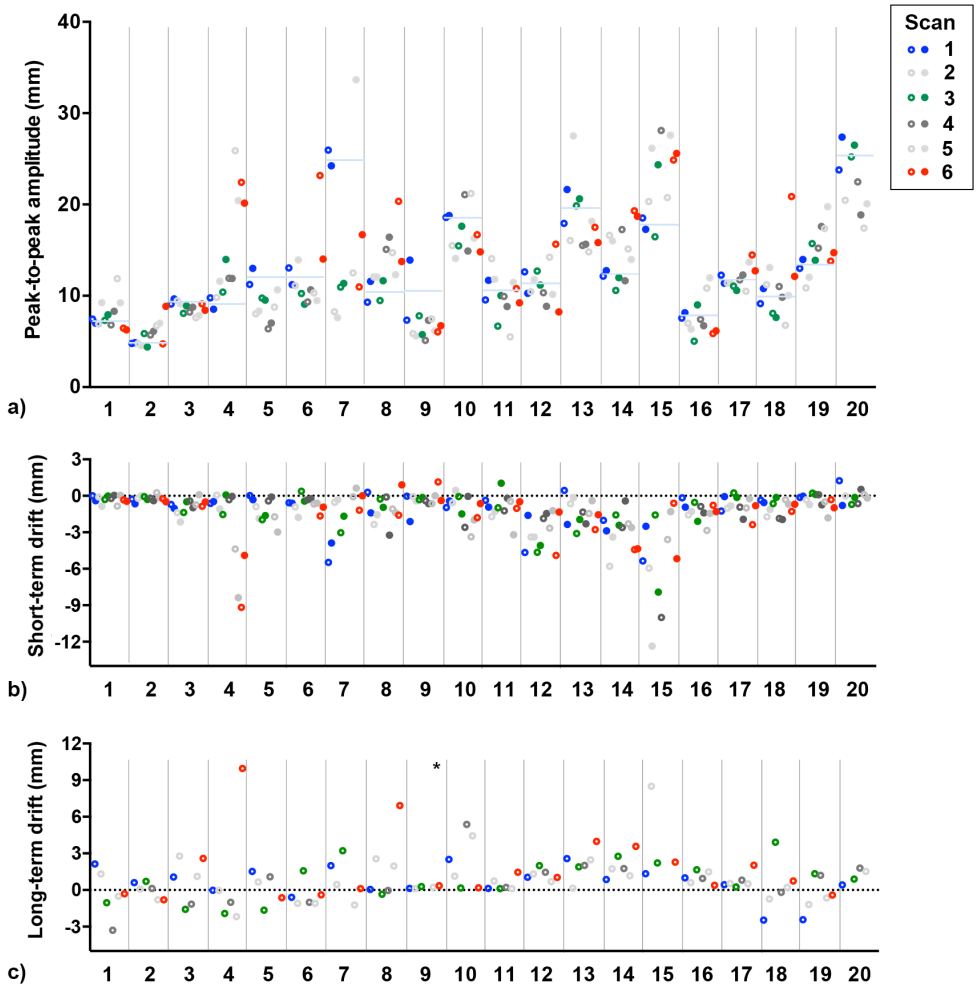


Figure 6.4: Intra-fraction motion for the six different time points visualized for each patient. For every time point the peak-to-peak CC-motion is shown for both scans within the session. The second series is indicated with a closed circle. All patients had a distal tumor, except for patients 1 and 2 (both middle third) and 18 to 20 (gastroesophageal junction). In graph **a)** the peak-to-peak motion is shown. The blue solid lines indicate the mean peak-to-peak amplitude based on the two baseline scans. The estimated short-term (∼45 seconds) and long-term drift (∼10 minutes) are shown in the **b)** and **c)**, respectively. In both graphs y=0 is indicated with a dotted line and represents the absence of drift.
**The cine-MRI series of time point 4 for this patient were excluded because of table movement that occurred between the cine-MRI series.*

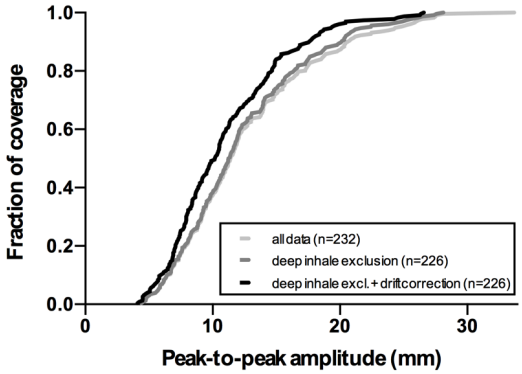


Figure 6.5: Fraction of scans with a peak-to-peak amplitude lower than the indicated peak-to-peak amplitude, both for the original dataset (n=232), the dataset after exclusion of deep inhales (n=226), and the dataset corrected for tumor drift (n=226). The shift to the left in the lines visualizes the decrease in average peak-to-peak amplitudes after drift correction.

Table 6.2: Cranio-caudal tumor motion

	All data (n=232)			Deep inhale (n=226)			Drift corrected (n=226)		
	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
p-t-p (mm)	12.7	5.6	4.4-33.7	12.3	5.2	4.4-28.1	11.0	4.6	4.1-26.6
C95 (mm)	10.9	4.8	3.5-27.3	10.6	4.6	3.5-26.7	9.8	4.0	3.9-25.4

p-t-p = peak-to-peak.

Discussion

In this study, individual intra-fraction motion of esophageal tumors was weekly monitored during the course of nCRT. Interestingly, it was shown that the motion amplitude varied considerably between and within patients. Furthermore, it was found that a baseline scan is not representative to predict tumor motion throughout treatment. Similarly to other studies, it was found that the largest tumor motion occurred in the CC direction.

The large variations in CC amplitudes could be explained by tumor drift within cine-MRI series (short-term drift), but also by a more constant tumor motion pattern that includes a deep inhale, resulting in large peak-to-peak amplitudes. The duration of one fraction of 1.8 Gy (approximately 2 minutes) is comparable with the length of one cine-MRI series, which makes the tumor motion found in the cine-MRI series representative for motion that could occur during the delivery of one fraction. Furthermore, long-term drift that occurred within a 10-minute time-interval was quantified, since

this could be representative for tumor drift in a time-interval needed for daily dose optimization of the dose plan between positioning of the patient and the actual start of radiation delivery. The radiation dose distributions of patients with large tumor motion have been shown to be influenced by this tumor motion [163, 164]. Future work will concentrate on the possible dosimetric consequences of tumor motion for radiotherapy delivery in esophageal cancer and the potential of using personalized motion margins. Furthermore, strategies to mitigate the effect of tumor motion deserve further investigation.

Some previous studies have mentioned pre-treatment identification of regular and irregular breathers [165]. A patient-specific tumor margin for the delivery of radiotherapy could then be determined based on 1 or 2 initial cine-MRI series. However, this is contradicted by the finding of the current study, since large variation in intra-fraction motion is observed in some patients, indicating that an estimation of tumor motion prior to the start of treatment may not represent the actual tumor motion during treatment. The application of fixed margins for esophageal tumors during radiotherapy, e.g. using the well known margin recipe [166], might not be sufficient when taking the complexity and large variety of tumor motion found in this study into account. However, knowledge of the tumor motion during radiotherapy could enable more conformal dose delivery and individualized treatment strategies. Currently available systems in which tumor tracking during dose delivery is possible are the CyberKnife, using X-ray imaging [52], and the Calypso, using several internal markers for tumor tracking [53]. Furthermore, an integrated MRI and Linac provides the advantages of an MRI system, enabling precise tracking of both tumor and surrounding soft tissues during dose delivery [54–57]. Ultimately, motion margins and consequently unnecessary radiation dose to surrounding organs at risk can be reduced by these approaches.

The conclusion of this study differs from that of Jin et al. [167], who state that motion defined in pre-treatment 4DCT is representative for motion throughout treatment. Furthermore, they stated that 97% of the cases had a motion amplitude of less than 10 mm. While we observed that, after drift correction, only 49% showed a peak-to-peak amplitude of less than 10 mm. Several differences in study setup can explain this. Firstly, both studies report on relatively small patient groups. In our study all patients were male and mainly distal tumors were included, while the study of Jin et al. includes several females and shows a larger variety in tumor locations. Especially the more proximally located tumors exhibit the smallest motion [48] and variability [167], which explains the smaller overall reported variability. Furthermore, tumor motion between males and females could also differ, based on anatomical variations of the thorax or a different breathing pattern [168, 169]. Secondly, 4D(CB)CT reconstruction might obscure tumor drift and large inhales, which could be the origin of the larger variability in intra-fraction motion observed in some patients in our

MR-based study.

Some limitations apply to our study. First, our study uses temporal 2D datasets, which could lead to the presence of some out-of-plane motion. The analysis was performed on a coronal slices, so out-of-plane motion would be in the AP direction. The average motion observed in this plane is 3.8 mm, which is below the scanned slice thickness of 5 mm. However, the difference in CC motion found between the coronal and sagittal scans hints to the presence of some out-of-plane motion, which could result in an underestimation or overestimation of tumor motion. Second, cine-MRI series were acquired with both arms positioned next to the body, this could have influenced the reported tumor motion amplitude when compared to the currently used radiotherapy position where both arms are raised above the head. However, the body positioning was consequent throughout all cine-MRI series. Furthermore, this positioning is not expected to influence the main findings of the current study regarding the non-representativeness of a baseline scan for tumor motion throughout treatment and varying motion amplitudes between and within patients.

The size of the motion found in 20 esophageal tumors in 3 different directions presented in this study compares well with previously published work in esophageal cancer [48, 46]. Previous published work also described correlations between clinical tumor stage and the mobility of the esophageal tumor. A significantly smaller tumor motion was found in cT4 tumors compared to cT1-T3 tumors, possibly due to the fact that cT4 tumors have an increased adhesion to surrounding tissues [49]. In the current study no statistical analyses were performed to assess the association between tumor location [48] or clinical tumor stage [49] and tumor motion due to the relatively small sample size including mainly distal tumors. In a larger, more heterogeneous cohort, the relation and clinical characteristics could be studied more extensively.

In conclusion, the largest esophageal tumor motion prior to, and throughout the course of nCRT, was found in CC direction, followed by the AP and LR direction using cine-MRI series. Tumor motion in CC direction is highly variable between and within patients during the course of nCRT, as well as within one cine-MRI series (45 seconds) and over a 10-minute interval. It was shown that esophageal tumor motion does not only comprise of respiratory motion, but is also caused by a tumor drift in some patients. The stochastic nature of both deep inhales and tumor drift indicates that real-time tumor motion management during radiotherapy is a prerequisite to safely reduce treatment margins.

Chapter 7

Summary and general discussion

Personalizing treatment: differentiation based on response

Esophageal cancer is known to have low 5-year overall survival rates [11–14]. Furthermore, the degree of response is associated with patient prognosis, with the most favorable long-term prognosis for the 29% of patients reaching pathologic complete response [11, 14, 58, 60, 61]. For the introduction of personalized treatment schemes accurate prediction or assessment of response is required in which, depending on response, patients could follow a different treatment path. This could include for example a wait-and-see approach with omission of surgery, intensification of neoadjuvant therapy or primary surgery. An example of a randomized trial in which a wait-and-see approach is currently being investigated for patients with esophageal cancer is the SANO trial, in which suspected complete responders are allocated in a surgery or active surveillance arm [170]. This study uses endoscopy with bite-on-bite and ^{18}F-FDG-PET/CT at different moments in time to predict response.

In the preceding chapters a range of topics has been discussed with a main focus on using MR-based (functional) techniques to predict response to neoadjuvant treatment for patients with esophageal cancer. The potential of using quantitative dynamic contrast-enhanced (DCE-) MRI for response prediction to neoadjuvant chemoradiotherapy was assessed (**Chapter 2**). The association between the ^{18}F-FDG-PET/CT parameter SUV, which is widely used clinically, and the quantitative diffusion-weighted (DW-) MRI parameter ADC was investigated (**Chapter 3**) as well as the complementary value of combining DCE-MRI with DW-MRI for response prediction (**Chapter 4**). A methodical study on intravoxel incoherent motion (IVIM) is described in (**Chapter 5**). Finally, the variations in esophageal tumor motion during radiotherapy are reported (**Chapter 6**).

Dynamic contrast-enhanced MRI

Visualizing tissue perfusion with dynamic contrast-enhanced (DCE-) MRI, enables the differentiation of tumorous tissue from healthy tissue, as tumors are characterized by angiogenesis paired with increased vascular permeability [35, 36]. Additionally, in several other tumor sites it was shown that DCE-MRI can be used for response prediction [76, 77].

One of the most used tracer-kinetic models for analysis of DCE-MRI is the Tofts model [171], providing a link between the tumor physiology and the measured data. However, this model is highly dependent on the shape of the arterial input function (AIF) for which often a generic or population-specific AIF is used. A recent publication using the Tofts model concluded that measures from the Tofts model were found effective predictors for the efficacy of chemoradiotherapy in patients with

esophageal cancer [172]. However, another study showed limited reproducibility of the Tofts model in pancreatic ductal adenocarcinoma, with for most parameters a 'within subject coefficient of variation' in the order of 20% [173]. In our initial study including three MRI scans (prior to treatment, during treatment and before surgery) large variations in peak height were found both between and within patients. Therefore, in our analysis of DCE-MRI in **Chapter 2**, the area-under-the-concentration time curve (AUC) was calculated as quantitative parameter, since this more simple and robust approach is not influence by the AIF as input parameter. We found that changes in the AUC within the tumor throughout the first two to three weeks of treatment differed significantly between good (including also near-pCR) and poor responders, potentially allowing for early prediction of treatment response. In addition, the change in AUC prior to surgery compared to the baseline-AUC was found to differ significantly between pathologic complete responders and non-pathologic complete responders. In this chapter it was concluded that DCE-MRI showed potential for treatment response assessment and prediction of pathologic complete response and good response on chemoradiotherapy in patients with esophageal cancer.

Since the potential in relation to response prediction of this MR-modality is increasingly shown for esophageal cancer, we foresee that in the future DCE-MRI will be increasingly included in standard diagnostic MR-protocols since this could be a useful adjunct to the current available modalities. However, larger trials researching the optimal analysis approach and time point for response prediction remain of great essence.

Diffusion-weighted MRI

Tumorous tissue is known for its high cellularity, resulting in a restricted diffusion compared to healthy tissue, which can be measured with diffusion-weighted (DW-) MRI [37, 38, 121]. When neoadjuvant treatment is effective tumor cellularity is expected to decrease, hence the apparent diffusion coefficient (ADC) has been evaluated for its potential to function as a biomarker for treatment response [174, 175]. An earlier study in our group showed that a difference in ADC in the tumor between scans during treatment (2-3 weeks) and baseline scans was highly predictive for histopathologic response in patients with esophageal cancer [66]. These findings were confirmed in recent publications concerning esophageal cancer [172, 176]. Interestingly, it was shown that DW-MRI can provide important, additional information in relation to both T and N staging and initial treatment decision-making compared to e.g. CT or endoscopic ultrasonography, making this technique a useful adjunct in the clinic [177].

^{18}F-FDG-PET/CT & DW-MRI

Whole-body ^{18}F-fluorodeoxyglucose positron emission tomography with integrated computed tomography (^{18}F-FDG-PET/CT) is often used in a pre-treatment setting to detect distant metastases and regional lymphadenopathy [9, 89, 90] and also serves as a useful adjunct to conventional pre-treatment staging modalities for esophageal cancer [88]. In addition, ^{18}F-FDG-PET/CT has also shown potential for response prediction [23, 24].

In other tumor sides the potential of whether ADC as measured by DW-MRI and standardized uptake value (SUV) as measured by ^{18}F-FDG-PET/CT can provide complementary information has been a recent topic of research, reporting correlations between SUV and ADC parameters [178, 179]. At present, it is unclear whether ADC and SUV can provide complementary prognostic and predictive information in patients with esophageal cancer. Therefore, the correlation between ADC and SUV in patients with esophageal cancer was researched in **Chapter 3**. No correlations were found between tumor ADC and SUV, regardless of the underlying histological subtype. Also, no associations of ADC and/or SUV with tumor stage or grade were found. Therefore, it was concluded that tumor ADC and SUV values may have complementary roles as imaging markers in the prediction of survival and evaluation of response to treatment in esophageal cancer. In designing a multi-parametric multi-modality study to evaluate the efficiency of predicting the response of esophageal cancer to neoadjuvant treatment, we therefore see a place for both ^{18}F-FDG-PET/CT and DW-MRI.

DCE-MRI & DW-MRI

Chapter 4 is a continuation of both studies researching DCE-MRI and DW-MRI. In this chapter the complementary value was researched of both modalities. Since both modalities visualize different tumor characteristics, combining these modalities has been researched in multiple studies concerning other tumor sites [42, 43]. These studies reported superior descriptive accuracies of combined analyses compared to separate analyses. In this study, uni- and multivariable logistic regression models were used to predict response to nCRT in patients with esophageal cancer. It was found that multivariable analysis of DCE-MRI and DW-MRI showed complementary value reaching higher predictive values.

Earlier studies regarding different tumor sites also reported superior descriptive accuracies of combined analyses of DW-MRI and DCE-MRI compared to separate analyses [42, 43]. An important conclusion presented in a recent meta-analysis concerning breast cancer researching both DW-MRI and DCE-MRI for predicting response revealed that the type of MRI scan and the time between the start of neoadjuvant treatment and MRI significantly influenced the sensitivity and specificity [180]. Due

to a limited number of multiparametric studies available for breast cancer, no direct comparison between DW-MRI and DCE-MRI was presented.

However, one drawback found in this study was that the quality of the transverse DW-MR images was not stable, showing large breathing artifacts in some patients. Therefore, part of the patients had to be excluded from the analysis. As a result, several adjustments (details in **Chapter 5**) were made in the MRI scanning protocol which was then used for the REACT-trial (abbreviation for: REpeated magnetic resonance imaging in esophageal cancer for Adaptive radiation treatment planning during ChemoradioTherapy).
The REACT-trial included acquisition of scans at six time-points, including one baseline scans and weekly scans during 5 weeks of neoadjuvant treatment. The primary objective was to develop a patient-specific adaptive radiotherapy planning strategy using MRI with more precise target coverage and critical organ sparing by safely reducing treatment margins for resectable esophageal cancer irradiation. Furthermore, our previous studies acquired scans around the second week of treatment based on two meta-analyses showing this timepoint as most predictive for pathologic response [23, 24, 67]. **Chapters 2 and 4** indeed showed potential in the prediction of response using this time-point. However, since the optimal timing of scanning for response prediction is still unclear, the secondary objective was to find the optimal timing for MRI guided response assessment to predict pathological response to nCRT as determined after surgery. One of the main adjustments made to the scanning protocol included a change in scanning plane of the diffusion-weighted images from transverse to coronal, resulting in higher image quality. Since the largest occurring motion (breathing motion) appears in line with the scanning plane, most of the motion artifacts produced can be corrected for in additional post-processing [142]. Furthermore, the number of scanned b-values was increased to 13 enabling more advanced fitting methods for the signal decay, also known as intravoxel incoherent motion MRI.

Intravoxel incoherent motion MRI

In intravoxel incoherent motion (IVIM-) MRI a high number of diffusion-weighted MRI scans is acquired with a wide range of b-values. IVIM-MRI has been a topic of research for several reasons. It has the potential of increasing accuracy of response prediction compared to regular mono-exponential fitting of ADC values in DW-MRI, with its ability to isolate diffusion and perfusion components. Furthermore, since in IVIM-MRI also perfusion-related parameters are fitted, it might probe properties similar to DCE-MRI. Although this is based on different mechanisms, in case of high correlation, injection of intravenous contrast could become obsolete. Although the insertion of an i.v. line for contrast injection was only scored as causing most discomfort during the whole MR procedure in 4% [181], omitting DCE-MRI would pave the way of frequent MRI scans with reduced time needed in the clinical workflow

and scan time (scored most discomfort in 22%). On the other hand, it should be noted that IVIM protocols are known for their high noise levels, which scored highest on patient discomfort (26%).

So far, varying results have been reported for the correlation between the IVIM perfusion-related parameters and DCE-MRI parameters outside the esophagus, with some studies showing significant correlations [138,182] in contrast to others [183,184]. However, before researching the correlation between IVIM and other modalities and the potential of using IVIM in relation to response prediction [185], obtaining a reliable fitting procedure is of great importance.

Several studies compared different approaches for fitting the well-known Bihan model [39], and concluded that large differences can occur in fitting parameters D (diffusion), f (perfusion fraction) and D^* (pseudo-diffusion) between different approaches [40, 141]. Therefore in **Chapter 5**, as a first step towards researching the potential of IVIM-MRI for patients with esophageal cancer, the influence of using three different fitting approaches (Bihan-fixed, Bihan-free and Bayesian) was researched on voxel-level. Model performance was assessed by calculation of R^2 values, which calculates to what extend the model fit describes the IVIM datapoints. Although no large differences were found in model performance, it was found that differences in trend occurred between the different fitting approaches. It is expected that these differences influence the accuracy and precision in longitudinal studies for response assessment in which IVIM will be used. Therefore, it was concluded that it is of great importance to explicit report specifications of fitting approaches used in IVIM studies to be able to compare outcomes between studies.

Personalizing radiotherapy treatment: safety margins

Another field which serves potential for personalizing the current treatment strategies is researching the use of personalized tumor margins during radiotherapy. Real-time tracking of the tumor during dose delivery becomes reality with an integrated MRI and Linac [54–57] or systems like the CyberKnife (AccurayTM, Inc, USA) [52] or Calypso (Varian Medical Systems, Palo Alto, CA) [53]. As a first step towards real-time tumor tracking for esophageal cancer the variations in intra-fraction motion were researched on different time scales (**Chapter 6**). Patients treated with nCRT for esophageal cancer underwent six sequential MRI scans (including one baseline scan). To quantify intra-fraction motion, two cine-MRI series were acquired with approximately 10 minutes interval, both in coronal and sagittal plane. Tumor motion was assessed in cranio-caudal (CC), anterior-posterior (AP) and left-right (LR) direction. In the CC-plane largest breathing motion was found, therefore CC-motion patterns were further analyzed for the presence of deep inhales and tumor drift. This study

showed that intra-fraction tumor motion was found to be highly variable between and within patients with esophageal cancer over the course of nCRT treatment. Large variation in intra-fraction motion was observed in some patients, indicating that an estimation of tumor motion prior to the start of treatment may not represent the actual tumor motion during treatment. A strategy in which a patient-specific margin is determined based on 1 or 2 initial cine-MRI series does therefore not reflect the true tumor motion during treatment. So, pre-treatment identification of regular and irregular breathers [165] would not function as a proper solution. This study therefore concluded that the stochastic nature of tumor drift indicated that real-time tumor motion management during radiotherapy is a prerequisite to safely reduce treatment margins.

In addition to a possible reduction of the dose to organs at risk which could be achieved when reducing treatment margins, real-time tumor tracking could also enable dose painting. To increase the tumor local control rates while sparing healthy tissue, in dose painting the regular uniform dose is replaced by a individualized dose distribution [186]. This is designed such that an additional dose is delivered to subvolumes with for example high radioresistance. These regions could be identified by functional imaging in case response prediction is further exploited towards voxel-based response predictions. Several recent studies concerning different tumor sites have shown the (potential) benefit of using dose painting over regular dose description in relation to local tumor control and survival [187–189]. The high accuracy needed for this procedure underlines the necessity of tumor tracking during dose painting, and consequently the use of the sophisticated radiotherapy systems which are mentioned above.

Future perspectives

This thesis focuses mainly on the use of quantitative MRI analysis for the prediction of response. Although the studies presented in this thesis demonstrated the predictive value of several image-based techniques, a couple important questions remain to be further researched.
First, the conclusions presented in this thesis should be validated in larger patient cohorts, enabling the development of a robust and accurate prediction model. Second, the reproducibility of the different modalities/parameters should be researched since clinically valid use requires that measurement variations in a given patient be less than that observed by different observers or measurement. This reproducibility is two-fold, in which both the inter- and intra-observer variations in delineations contribute, as well as the reproducibility of the measurement itself. In our most recent studies, a semi-automatic method was used after which only minor manual adjustments were made based on tumor circumference changes through time. In the future

study including a larger patient cohort and scans with improved quality, a delineation study researching both the inter- and intra-observer variability and the sensibility for changes over time could offer new insides. Studies already published concerning esophageal cancer showed sufficient reproducibility comparing two measurements and two observers [190] and a higher reproducibility in semi-automated delineations techniques compared to manual delineations [92]. Furthermore, a recent publication showed promising results researching the repeatability, reproducibility and accuracy of both DW-MRI and T_1-mapping in the breast in healthy subjects and phantoms (ice-water and gel phantom, respectively) across three sites [191].

Beside the quantitative image-based techniques that were studied in this thesis clinical parameters and qualitative assessment of T_2-weighted scans or diffusion-weighted scans could provide additional information. In addition, several studies showed potential for the use of molecular biomarkers for response prediction in patients with esophageal cancer [68–70]. Likewise, the presence of circulating tumor cells (CTCs) in the blood is known to be associated with poor clinical outcome [192]. Since this method is minimally invasive, it has the potential to be a valuable tool in prediction of treatment outcome. However, this technique remains technically challenging since CTC enrichment involves isolating these very rare CTCs from billions of red blood cells and millions of leukocytes per milliliter of whole blood [193]. It was found that patients with esophageal cancer presenting with an initial negative CTC score, meaning that no circulating tumor cells were found, or a change from positive to negative showed a favorable prognosis concerning overall survival and progression free survival compared to patients with a positive CTC score both at baseline and after treatment. Pre- and intra-therapeutic CTC enrichment was found to show potential to e.g. improve preoperative staging [194, 195].

A multicenter trial recently initiated in our hospital will address several of the above mentioned points. The main objective of the PRIDE-trial (abbreviation for: Preoperative Image-guided Identification of Response to Neoadjuvant Chemoradiotherapy in Esophageal cancer) is to develop a multi-parametric prediction model that predicts the probability of pathologic complete response to nCRT in esophageal cancer by integrating DW-MRI and DCE-MRI in conjunction with ^{18}F-FDG-PET/CT scans acquired prior to, during and after administration of nCRT. Since this study will include a total of 200 patients validation of the conclusions from the preceding chapters in a larger cohort will be possible. In addition, new promising techniques are also being investigated, with one of the secondary objectives to evaluate the presence of, and changes in, CTCs in the blood during nCRT as a biomarker for a patients response to nCRT, the detection of residual disease after nCRT and disease-free and overall survival.

An ultimate goal would be a personalized treatment scheme in which margins can be reduced to reduce toxicity to surrounding organs at risk and dose painting is

possible due to real-time tracking of the tumor and accurate, voxel-based, response prediction. Additionally, these non- and/or minimally invasive response prediction techniques would enable a wait-and-see approach in which surgery is only performed in patients who do not reach a pathologic complete response.

In conclusion, the studies presented in this thesis showed the potential of functional MRI and PET in the prediction of response in patients with esophageal cancer undergoing neoadjuvant chemoradiotherapy. Furthermore, it was found that different modalities (e.g. DCE-MRI, DW-MRI and ^{18}F-FDG-PET/CT) could be of complementary value in predicting response, each visualizing different physiological processes in the tumor. Therefore, the use of multi-parametric multimodality models should be further researched in larger patient cohorts, in which the emerging field of machine learning/deep learning could be profitable in developing response prediction models. Publications on the potential of these models in relation to biology and medicine are emerging [196–198], with a recent publication concerning esophageal cancer showing the value of T_2W-MRI radiomic analysis combined with supervised machine-learning algorithms for response prediction [199]. In addition, it was shown that large variations were found in tumor motion both between and within patients with esophageal cancer and the stochastic nature of the variations in motion demonstrated the importance of real-time tumor motion management during radiotherapy in order to safely reduce treatment margins.

Chapter 8

References

References

[1] IKNL (2018). Nederlandse kankerregistratie; http://www.cijfersoverkanker.nl.

[2] Torre, L.a., Bray, F., Siegel, R.L., et al. (2015). Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, volume 65(2):87–108. ISSN 00079235.

[3] Enzinger, P.C. and Mayer, R.J. (2003). Esophageal Cancer. *New England Journal of Medicine*, volume 349(23):2241–2252. ISSN 0028-4793.

[4] Pennathur, A., Gibson, M.K., Jobe, B.a., and Luketich, J.D. (2013). Oesophageal carcinoma. *The Lancet*, volume 381(9864):400–412. ISSN 01406736.

[5] Arnal, M.J.D., Arenas, Á.F., and Arbeloa, Á.L. (2015). Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries. *World Journal of Gastroenterology*, volume 21(26):7933–7943. ISSN 22192840.

[6] Dikken, J.L., Lemmens, V.E., Wouters, M.W.J.M., et al. (2012). Increased incidence and survival for oesophageal cancer but not for gastric cardia cancer in the Netherlands. *European Journal of Cancer*, volume 48(11):1624–1632. ISSN 09598049.

[7] Giovagnoni, a., Valeri, G., and Ferrara, C. (2002). MRI of esophageal cancer. *Abdominal Imaging*, volume 27(4):361–366. ISSN 0942-8925.

[8] Wallace, M.B., Nietert, P.J., Earle, C., et al. (2002). An analysis of multiple staging management strategies for carcinoma of the esophagus: Computed tomography, endoscopic ultrasound, positron emission tomography, and thoracoscopy/laparoscopy. *Annals of Thoracic Surgery*, volume 74(4):1026–1032. ISSN 00034975.

[9] Kato, H., Miyazaki, T., Nakajima, M., et al. (2005). The incremental effect of positron emission tomography on diagnostic accuracy in the initial staging of esophageal carcinoma. *Cancer*, volume 103(1):148–156. ISSN 0008-543X.

[10] van Rossum, P.S.N., van Lier, A.L.H.M.W., Lips, I.M., et al. (2014). Imaging of oesophageal cancer with FDG-PET/CT and MRI. *Clinical radiology*. ISSN 1365-229X.

[11] van Hagen, P., Hulshof, M.C.C.M., van Lanschot, J.J.B., et al. (2012). Preoperative chemoradiotherapy for esophageal or junctional cancer. *The New England journal of medicine*, volume 366(22):2074–84. ISSN 1533-4406.

[12] Omloo, J.M.T., Lagarde, S.M., Hulscher, J.B.F., et al. (2007). Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the mid/distal esophagus: five-year survival of a randomized clinical trial. *Annals of surgery*, volume 246(6):992–1001. ISSN 0003-4932.

[13] Sjoquist, K.M., Burmeister, B.H., Smithers, B.M., et al. (2011). Survival after neoadjuvant chemotherapy or chemoradiotherapy for resectable oesophageal carcinoma: an updated meta-analysis. *The Lancet. Oncology*, volume 12(7):681–92. ISSN 1474-5488.

[14] Shapiro, J., van Lanschot, J.J.B., Hulshof, M.C.C.M., et al. (2015). Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): Long-term results of a randomised controlled trial. *The Lancet Oncology*, volume 16(9):1090–1098. ISSN 14745488.

[15] Noordman, B.J., Verdam, M.G., Lagarde, S.M., et al. (2018). Effect of Neoadjuvant Chemoradiotherapy on Health-Related Quality of Life in Esophageal or Junctional Cancer: Results From the Randomized CROSS Trial. *Journal of Clinical Oncology*, volume 36(3):268–275. ISSN 0732-183X.

[16] Goense, L., van Rossum, P.S.N., Tromp, M., et al. (2016). Intraoperative and postoperative risk factors for anastomotic leakage and pneumonia after esophagectomy for cancer. *Diseases of the Esophagus*, (February):1–10. ISSN 11208694.

[17] Mandard, A., Dalibard, F., and Mandard, J. (1994). Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer*, pp. 2680–2686.

[18] Wani, S., Early, D., Edmundowicz, S., and Sharma, P. (2012). Management of High-Grade Dysplasia and Intramucosal Adenocarcinoma in Barrett's Esophagus. *Clinical Gastroenterology and Hepatology*, volume 10(7):704–711. ISSN 15423565.

[19] Goense, L., van Rossum, P.S.N., Kandioler, D., et al. (2016). Stage-directed individualized therapy in esophageal cancer. *Annals of the New York Academy of Sciences*, volume 1381(1):50–65. ISSN 17496632.

[20] Napier, K.J. (2014). Esophageal cancer: A Review of epidemiology, pathogenesis, staging workup and treatment modalities. *World Journal of Gastrointestinal Oncology*, volume 6(5):112. ISSN 1948-5204.

[21] Lin, J.L. (2013). T1 esophageal cancer, request an endoscopic mucosal resection (EMR) for in-depth review. *Journal of Thoracic Disease*, volume 5(3):353–356. ISSN 20721439.

[22] Wani, S., Drahos, J., Cook, M.B., et al. (2014). Comparison of endoscopic therapies and surgical resection in patients with early esophageal cancer: A population-based study. *Gastrointestinal Endoscopy*, volume 79(2):224–232.e1. ISSN 00165107.

[23] Kwee, R. (2010). Prediction of Tumor Response to Neoadjuvant Therapy in Patients with Esophageal Cancer with Use of 18 F FDG PET: A Systematic Review. *Radiology*, volume 254(3):707–717.

[24] Chen, Y.m., Pan, X.f., Tong, L.j., Shi, Y.p., and Chen, T. (2011). Can F-fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. *Nuclear medicine communications*, volume 32(11):1005–10. ISSN 1473-5628.

[25] Khoo, V.S. and Joon, D.L. (2006). New developments in MRI for target volume delineation in radiotherapy. *The British Journal of Radiology*, volume 79(special_issue_1):S2–S15. ISSN 0007-1285.

[26] Hentschel, B., Oehler, W., Strauß, D., Ulrich, A., and Malich, A. (2011). Definition of the CTV Prostate in CT and MRI by Using CTMRI Image Fusion in IMRT Planning for Prostate Cancer. *Strahlentherapie und Onkologie*, volume 187(3):183–190. ISSN 0179-7158.

[27] van Rossum, P.S.N., van Hillegersberg, R., Lever, F.M., et al. (2013). Imaging strategies in the management of oesophageal cancer: what's the role of MRI? *European Radiology*, volume 23(7):1753–1765. ISSN 0938-7994.

[28] Dappa, E., Elger, T., Hasenburg, A., et al. (2017). The value of advanced MRI techniques in the assessment of cervical cancer: a review. *Insights into Imaging*, volume 8(5):471–481. ISSN 18694101.

[29] Hanahan, D. and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, volume 144(5):646–674. ISSN 00928674.

[30] Padhani, A.R. (2002). Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. *Journal of magnetic resonance imaging*, volume 16(4):407–22. ISSN 1053-1807.

[31] Charles-Edwards, E.M. and DeSouza, N.M. (2006). Diffusion-weighted magnetic resonance imaging and its application to cancer. *Cancer imaging : the official publication of the International Cancer Imaging Society*, volume 6(1):135–43. ISSN 1470-7330.

[32] Oberholzer, K., Pohlmann, A., Schreiber, W., et al. (2008). Assessment of tumor microcirculation with dynamic contrast-enhanced MRI in patients with esophageal cancer: Initial experience. *Journal of Magnetic Resonance Imaging*, volume 27(6):1296–1301. ISSN 10531807.

[33] Aoyagi, T., Shuto, K., Okazumi, S., et al. (2011). Apparent diffusion coefficient values measured by diffusion-weighted imaging predict chemoradiotherapeutic effect for advanced esophageal cancer. *Digestive Surgery*, volume 28(4):252–257. ISSN 02534886.

[34] Imanishi, S., Shuto, K., Aoyagi, T., et al. (2013). Diffusion-weighted magnetic resonance imaging for predicting and detecting the early response to chemoradiotherapy of advanced esophageal squamous cell carcinoma. *Digestive surgery*, volume 30(3):240–8. ISSN 1421-9883.

[35] Carmeliet, P. and Jain, R.K. (2000). Angiogenesis in cancer and other diseases. *Nature*, volume 407:249–257. ISSN 0028-0836.

[36] Khalifa, F., Soliman, A., El-Baz, A., et al. (2014). Models and methods for analyzing DCE-MRI: A review. *Medical Physics*, volume 41(12):124301. ISSN 0094-2405.

[37] Thoeny, H.C. and Ross, B.D. (2010). Predicting and monitoring cancer treatment response with diffusion-weighted MRI. *Journal of Magnetic Resonance Imaging*, volume 32(1):2–16. ISSN 10531807.

[38] Padhani, A.R., Liu, G., Koh, D.M., et al. (2009). Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia*, volume 11(2):102–125. ISSN 1476-5586 (Electronic) 1476-5586 (Linking).

[39] Le Bihan, D., Breton, E., Lallemand, D., et al. (1988). Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*, volume 168(2):497–505. ISSN 0033-8419.

[40] Barbieri, S., Donati, O.F., Froehlich, J.M., and Thoeny, H.C. (2016). Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs. *Magnetic Resonance in Medicine*, volume 75(5):2175–2184. ISSN 15222594.

[41] Cho, G.Y., Moy, L., Zhang, J.L., et al. (2015). Comparison of fitting methods and b-value sampling strategies for intravoxel incoherent motion in breast cancer. *Magnetic Resonance in Medicine*, volume 74(4):1077–1085. ISSN 15222594.

[42] Li, X., Abramson, R.G., Arlinghaus, L.R., et al. (2015). Combined DCE-MRI and DW-MRI for Predicting Breast Cancer Pathological Response After the First Cycle of Neoadjuvant Chemotherapy. *Investigative Radiology*, volume 50(4):195–204. ISSN 0020-9996.

[43] Chawla, S. and Kim, S. (2013). Pretreatment diffusion-weighted and dynamic contrast-enhanced MRI for prediction of local treatment response in squamous cell carcinomas of the head and neck. *American Journal of Roentgenology*, volume 200(1):35–43.

[44] Intven, M., Monninkhof, E.M., Reerink, O., and Philippens, M.E.P. (2015). Combined T2w volumetry, DW-MRI and DCE-MRI for response assessment after neo-adjuvant chemoradiation in locally advanced rectal cancer. *Acta Oncologica*, (February 2016):1–8. ISSN 1651-226X (Electronic).

[45] Covello, M., Cavaliere, C., Aiello, M., et al. (2015). Simultaneous PET/MR head-neck cancer imaging: Preliminary clinical experience and multiparametric evaluation. *European Journal of Radiology*, volume 84(7):1269–1276. ISSN 18727727.

[46] Jin, P., Hulshof, M.C., de Jong, R., et al. (2016). Quantification of respiration-induced esophageal tumor motion using fiducial markers and four-dimensional computed tomography. *Radiotherapy and Oncology*, volume 118(3):492–497. ISSN 01678140.

[47] Patel, A.A., Wolfgang, J.A., Niemierko, A., et al. (2009). Implications of Respiratory Motion as Measured by Four-Dimensional Computed Tomography for Radiation Treatment Planning of Esophageal Cancer. *International Journal of Radiation Oncology Biology Physics*, volume 74(1):290–296. ISSN 03603016.

[48] Lever, F.M., Lips, I.M., Crijns, S.P.M., et al. (2013). Quantification of Esophageal Tumor Motion on Cine-Magnetic Resonance Imaging. *International journal of radiation oncology, biology, physics*. ISSN 1879-355X.

[49] Kobayashi, Y., Myojin, M., Shimizu, S., and Hosokawa, M. (2016). Esophageal motion characteristics in thoracic esophageal cancer: Impact of clinical stage T4 versus stages T1-T3. *Advances in Radiation Oncology*, volume 1(4):222–229. ISSN 24521094.

[50] Van Herk, M. (2004). Errors and Margins in Radiotherapy. *Seminars in Radiation Oncology*, volume 14(1):52–64. ISSN 10534296.

[51] Barnett, G.C., West, C.M.L., Dunning, A.M., et al. (2009). Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. *Nature Reviews Cancer*, volume 9(2):134–142. ISSN 1474-175X.

[52] Coste-Manière, E., Olender, D., Kilby, W., and Schulz, R.A. (2005). Robotic whole body stereotactic radiosurgery: clinical advantages of the Cyberknife integrated system. *The international journal of medical robotics + computer assisted surgery : MRCAS*, volume 1(2):28–39. ISSN 1478596X.

[53] Bell, L.J., Eade, T., Kneebone, A., et al. (2017). Initial experience with intra-fraction motion monitoring using Calypso guided volumetric modulated arc therapy for definitive prostate cancer treatment. *Journal of Medical Radiation Sciences*, volume 64(1):25–34. ISSN 20513909.

[54] Lagendijk, J.J.W., Raaymakers, B.W., Raaijmakers, A.J.E., et al. (2008). MRI/linac integration. *Radiotherapy and Oncology*, volume 86(1):25–29. ISSN 01678140.

[55] Kontaxis, C., Bol, G.H., Stemkens, B., et al. (2017). Towards fast online intrafraction replanning for free-breathing stereotactic body radiation therapy with the MR-linac. *Physics in Medicine and Biology*. ISSN 0031-9155.

[56] Mutic, S. and Dempsey, J.F. (2014). The ViewRay System: Magnetic Resonance-Guided and Controlled Radiotherapy. *Seminars in Radiation Oncology*, volume 24(3):196–199. ISSN 15329461.

[57] Kontaxis, C., Bol, G.H., Lagendijk, J.J.W., and Raaymakers, B.W. (2015). A new methodology for inter- and intrafraction plan adaptation for the MR-linac. *Physics in Medicine and Biology*, volume 60(19):7485–7497. ISSN 0031-9155.

[58] Donahue, J.M., Nichols, F.C., Li, Z., et al. (2009). Complete pathologic response after neoadjuvant chemoradiotherapy for esophageal cancer is associated with enhanced survival. *The Annals of thoracic surgery*, volume 87(2):392–9. ISSN 1552-6259.

[59] Allum, W.H., Stenning, S.P., Bancewicz, J., Clark, P.I., and Langley, R.E. (2009). Long-Term Results of a Randomized Trial of Surgery With or Without Preoperative Chemotherapy in Esophageal Cancer. *Journal of Clinical Oncology*, volume 27(30):5062–5067. ISSN 0732-183X.

[60] Chirieac, L.R., Swisher, S.G., Ajani, J.a., et al. (2005). Posttherapy pathologic stage predicts survival in patients with esophageal carcinoma receiving preoperative chemoradiation. *Cancer*, volume 103(7):1347–55. ISSN 0008-543X.

[61] Berger, A.C., Farma, J., Scott, W.J., et al. (2005). Complete response to neoadjuvant chemoradiotherapy in esophageal carcinoma is associated with significantly improved survival. *Journal of Clinical Oncology*, volume 23(19):4330–4337. ISSN 0732183X.

[62] Blazeby, J.M., Sanford, E., Falk, S.J., Alderson, D., and Donovan, J.L. (2005). Health-related quality of life during neoadjuvant treatment and surgery for localized esophageal carcinoma. *Cancer*, volume 103(9):1791–9. ISSN 0008-543X.

[63] Brücher, B.L.D.M., Weber, W., Bauer, M., et al. (2001). Neoadjuvant Therapy of Esophageal Squamous Cell Carcinoma : Response Evaluation by Positron Emission Tomography. *Annals of surgery*, volume 233(3):300–309.

[64] Schneider, P.M., Metzger, R., Schaefer, H., et al. (2008). Response evaluation by endoscopy, rebiopsy, and endoscopic ultrasound does not accurately predict histopathologic regression after neoadjuvant chemoradiation for esophageal cancer. *Annals of surgery*, volume 248(6):902–908. ISSN 0003-4932.

[65] Ajani, J.a., Correa, a.M., Hofstetter, W.L., et al. (2012). Clinical parameters model for predicting pathologic complete response following preoperative chemoradiation in patients with esophageal cancer. *Annals of Oncology*, volume 23(10):2638–2642. ISSN 09237534.

[66] van Rossum, P.S., van Lier, A.L., van Vulpen, M., et al. (2015). Diffusion-weighted magnetic resonance imaging for the prediction of pathologic response to neoadjuvant chemoradiotherapy in esophageal cancer. *Radiotherapy and Oncology*, volume 115(2):163–170. ISSN 01678140.

[67] Wieder, H.a., Ott, K., Lordick, F., et al. (2007). Prediction of tumor response by FDG-PET: comparison of the accuracy of single and sequential studies in patients with adenocarcinomas of the esophagogastric junction. *European journal of nuclear medicine and molecular imaging*, volume 34(12):1925–32. ISSN 1619-7070.

[68] Toxopeus, E.L.A., Nieboer, D., Shapiro, J., et al. (2015). Nomogram for predicting pathologically complete response after neoadjuvant chemoradiotherapy for oesophageal cancer. *Radiotherapy and Oncology*, volume 115(3):392–398. ISSN 18790887.

[69] Honing, J., Pavlov, K.V., Mul, V.E.M., et al. (2015). CD44, SHH and SOX2 as novel biomarkers in esophageal cancer patients treated with neoadjuvant chemoradiotherapy. *Radiotherapy and Oncology*, volume 117(1):152–8. ISSN 1879-0887.

[70] van Olphen, S.H., Biermann, K., Shapiro, J., et al. (2016). P53 and SOX2 Protein Expression Predicts Esophageal Adenocarcinoma in Response to Neoadjuvant Chemoradiotherapy. *Annals of surgery*. ISSN 1528-1140.

[71] Chang, E.Y., Li, X., Jerosch-Herold, M., et al. (2008). The Evaluation of Esophageal Adenocarcinoma Using Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Journal of Gastrointestinal Surgery*, volume 12(1):166–175. ISSN 1091-255X.

[72] Heethuis, S., van Rossum, P., Lips, I., et al. (2014). Comparison of image registration strategies to improve DCE-MRI uptake curves in esophageal cancer. In *56th annual meeting of the AAPM 2014 Austin, Texas, US. WE-G-18C-4*.

[73] Kanal, E., Maravilla, K., and Rowley, H.A. (2014). Gadolinium Contrast Agents for CNS Imaging: Current Concepts and Clinical Evidence. *American Journal of Neuroradiology*, volume 35(12):2215–2226. ISSN 0195-6108.

[74] Joo, I., Lee, J.M., Han, J.K., et al. (2014). Dynamic contrast-enhanced MRI of gastric cancer: Correlation of the perfusion parameters with pathological prognostic factors. *Journal of magnetic resonance imaging*. ISSN 1522-2586.

[75] Evelhoch, J.L. (2004). Magnetic Resonance Imaging Measurements of the Response of Murine and Human Tumors to the Vascular-Targeting Agent ZD6126. *Clinical Cancer Research*, volume 10(11):3650–3657. ISSN 1078-0432.

[76] Yuh, W.T.C., Mayr, N.A., Jarjoura, D., et al. (2009). Predicting Control of Primary Tumor and Survival by DCE MRI During Early Therapy in Cervical Cancer. *Investigative radiology*, volume 44(6):343–350.

[77] Li, X., Arlinghaus, L.R., Ayers, G.D., et al. (2014). DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: Pilot study findings. *Magnetic Resonance in Medicine*, volume 71(4):1592–1602. ISSN 07403194.

[78] Hermann, R.M., Horstmann, O., Haller, F., et al. (2006). Histomorphological tumor regression grading of esophageal carcinoma after neoadjuvant radiochemotherapy: which score to use? *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / I.S.D.E*, volume 19(5):329–34. ISSN 1120-8694.

[79] Vincent, J., Mariette, C., Pezet, D., et al. (2015). Early surgery for failure after chemoradiation in operable thoracic oesophageal cancer. Analysis of the non-randomised patients in FFCD 9102 phase III trial: Chemoradiation followed by surgery versus chemoradiation alone. *European Journal of Cancer*, volume 51(13):1683–1693. ISSN 18790852.

[80] Habr-Gama, A., Perez, R.O., Proscurshim, I., et al. (2008). Interval between surgery and neoadjuvant chemoradiation therapy for distal rectal cancer: does delayed surgery have an impact on outcome? *International journal of radiation oncology, biology, physics*, volume 71(4):1181–8. ISSN 0360-3016.

[81] Habr-Gama, A., Perez, R.O., Proscurshim, I., et al. (2006). Patterns of Failure and Survival for Nonoperative Treatment of Stage c0 Distal Rectal Cancer Following Neoadjuvant Chemoradiation Therapy. *Journal of Gastrointestinal Surgery*, volume 10(10):1319–1329. ISSN 1091255X.

[82] Parker, G.J., Roberts, C., Macdonald, A., et al. (2006). Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magnetic Resonance in Medicine*, volume 56(5):993–1000. ISSN 07403194.

[83] Li, S.P. and Padhani, A.R. (2012). Tumor response assessments with diffusion and perfusion MRI. *Journal of magnetic resonance imaging*, volume 35(4):745–63. ISSN 1522-2586.

[84] Roberts, C., Issa, B., Stone, A., et al. (2006). Comparative study into the robustness of compartmental modeling and model-free analysis in DCE-MRI studies. *Journal of magnetic resonance imaging*, volume 23(4):554–63. ISSN 1053-1807.

[85] Huang, W., Li, X., Chen, Y., et al. (2014). Variations of Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Evaluation of Breast Cancer Therapy Response: A Multicenter Data Analysis Challenge. *Translational Oncology*, volume 7(1):153–166. ISSN 19365233.

[86] Gwynne, S., Hurt, C., Evans, M., et al. (2011). Definitive Chemoradiation for Oesophageal Cancer - a Standard of Care in Patients with Non-metastatic Oesophageal Cancer. *Clinical Oncology*, volume 23(3):182–188. ISSN 09366555.

[87] Teoh, A.Y.B., Chiu, P.W.Y., Yeung, W.K., et al. (2013). Long-term survival outcomes after definitive chemoradiation versus surgery in patients with resectable squamous carcinoma of the esophagus: Results from a randomized controlled trial. *Annals of Oncology*, volume 24(1):165–171. ISSN 09237534.

[88] Wong, R., Walker-Dilks, C., and Raifu, A. (2012). Evidence-based Guideline Recommendations on the use of Positron Emission Tomography Imaging in Oesophageal Cancer. *Clinical Oncology*, volume 24(2):86–104. ISSN 09366555.

[89] Bar-Shalom, R., Guralnik, L., Tsalic, M., et al. (2005). The additional value of PET/CT over PET in FDG imaging of oesophageal cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, volume 32(8):918–924. ISSN 16197070.

[90] Van Westreenen, H.L., Westerterp, M., Bossuyt, P.M., et al. (2004). Systematic review of the staging performance of 18F- fluorodeoxyglucose positron emission tomography in esophageal cancer. *Journal of Clinical Oncology*, volume 22(18):3805–3812. ISSN 0732183X.

[91] Van Rossum, P.S.N., Van Lier, A.L.H.M.W., Lips, I.M., et al. (2015). Imaging of oesophageal cancer with FDG-PET/CT and MRI. *Clinical Radiology*, volume 70(1):81–95. ISSN 1365229X.

[92] Kwee, R.M., Dik, A.K., Sosef, M.N., et al. (2014). Interobserver Reproducibility of Diffusion-Weighted MRI in Monitoring Tumor Response to Neoadjuvant Therapy in Esophageal Cancer. *PLoS ONE*, volume 9(4).

[93] Sakurada, A., Takahara, T., Kwee, T.C., et al. (2009). Diagnostic performance of diffusion-weighted magnetic resonance imaging in esophageal cancer. *European Radiology*, volume 19(6):1461–1469. ISSN 0938-7994.

[94] Aoyagi, T. (2010). Evaluation of the clinical staging of esophageal cancer by using diffusion-weighted imaging. *Experimental and Therapeutic Medicine*, pp. 847–851. ISSN 1792-0981.

[95] Joye, I., Deroose, C.M., Vandecaveye, V., and Haustermans, K. (2014). The role of diffusion-weighted MRI and 18F-FDG PET/CT in the prediction of pathologic complete response after radiochemotherapy for rectal cancer: A systematic review. *Radiotherapy and Oncology*, volume 113(2):158–165. ISSN 01678140.

[96] Pickles, M.D., Gibbs, P., Lowry, M., and Turnbull, L.W. (2006). Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer. *Magnetic Resonance Imaging*, volume 24(7):843–847. ISSN 0730725X.

[97] Vandecaveye, V., Dirix, P., De Keyzer, F., et al. (2012). Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. *International Journal of Radiation Oncology Biology Physics*, volume 82(3):1098–1107. ISSN 03603016.

[98] Yushkevich, P.A., Piven, J., Hazlett, H.C., et al. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, volume 31(3):1116–1128. ISSN 10538119.

[99] Werner-Wasik, M., Nelson, A.D., Choi, W., et al. (2012). What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. *International Journal of Radiation Oncology Biology Physics*, volume 82(3):1164–1171. ISSN 03603016.

[100] Mukaka, M.M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, volume 24(3):69–71. ISSN 19957262.

[101] Bland, J. and Altman, D. (1986). Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *Lancet*, volume 327:307–310. ISSN 01406736.

[102] Hölscher, A.H., Drebber, U., Schmidt, H., and Bollschweiler, E. (2014). Prognostic Classification of Histopathologic Response to Neoadjuvant Therapy in Esophageal Adenocarcinoma. *Annals of Surgery*, volume 260(5):779–785. ISSN 0003-4932.

[103] Van Rossum, P.S., Goense, L., Meziani, J., et al. (2016). Endoscopic biopsy and EUS for the detection of pathologic complete response after neoadjuvant chemoradiotherapy in esophageal cancer: A systematic review and meta-analysis. *Gastrointestinal Endoscopy*, volume 83(5):866–879. ISSN 10976779.

[104] van Rossum, P.S.N., Fried, D.V., Zhang, L., et al. (2016). The Incremental Value of Subjective and Quantitative Assessment of 18F-FDG PET for the Prediction of Pathologic Complete Response to Preoperative Chemoradiotherapy in Esophageal Cancer. *Journal of Nuclear Medicine*, volume 57(5):691–700. ISSN 0161-5505.

[105] Surov, A., Stumpp, P., Meyer, H.J., et al. (2016). Simultaneous 18F-FDG-PET/MRI: Associations between diffusion, glucose metabolism and histopathological parameters in patients with head and neck squamous cell carcinoma. *Oral Oncology*, volume 58:14–20. ISSN 18790593.

[106] de Jong, A., Kwee, T.C., de Klerk, J.M., et al. (2014). Relationship between pretreatment {FDG-PET} and diffusion-weighted {MRI} biomarkers in diffuse large B-cell lymphoma. *Am J Nucl Med Mol Imaging*, volume 4(3):231–238.

[107] Fruehwald-Pallamar, J., Czerny, C., Mayerhoefer, M.E., et al. (2011). Functional imaging in head and neck squamous cell carcinoma: Correlation of PET/CT and diffusion-weighted imaging at 3 Tesla. *European Journal of Nuclear Medicine and Molecular Imaging*, volume 38(6):1009–1019. ISSN 16197070.

[108] Varoquaux, A., Rager, O., Lovblad, K.O., et al. (2013). Functional imaging of head and neck squamous cell carcinoma with diffusion-weighted MRI and FDG PET/CT: quantitative analysis of ADC and SUV. *European Journal of Nuclear Medicine and Molecular Imaging*, volume 40(6):842–852. ISSN 1619-7070.

[109] Wu, X., Korkola, P., Pertovaara, H., et al. (2011). No correlation between glucose metabolism and apparent diffusion coefficient in diffuse large B-cell lymphoma: A PET/CT and DW-MRI study. *European Journal of Radiology*, volume 79(2). ISSN 0720048X.

[110] Gu, J., Khong, P.L., Wang, S., et al. (2011). Quantitative Assessment of Diffusion-Weighted MR Imaging in Patients with Primary Rectal Cancer: Correlation with FDG-PET/CT. *Molecular Imaging and Biology*, volume 13(5):1020–1028. ISSN 1536-1632.

[111] Sakane, M., Tatsumi, M., Kim, T., et al. (2015). Correlation between apparent diffusion coefficients on diffusion-weighted MRI and standardized uptake value on FDG-PET/CT in pancreatic adenocarcinoma. *Acta Radiol*, volume 56(9):1034–1041. ISSN 0284-1851.

[112] Baba, S., Isoda, T., Maruoka, Y., et al. (2014). Diagnostic and Prognostic Value of Pretreatment SUV in 18F-FDG/PET in Breast Cancer: Comparison with Apparent Diffusion Coefficient from Diffusion-Weighted MR Imaging. *Journal of Nuclear Medicine*, volume 55(5):736–742. ISSN 0161-5505.

[113] Schwenzer, N.F., Schmidt, H., Gatidis, S., et al. (2014). Measurement of apparent diffusion coefficient with simultaneous MR/positron emission tomography in patients with peritoneal carcinomatosis: Comparison with 18F-FDG-PET. *Journal of Magnetic Resonance Imaging*, volume 40(5):1121–1128. ISSN 15222586.

[114] Wu, X., Pertovaara, H., Korkola, P., et al. (2014). Correlations between functional imaging markers derived from PET/CT and diffusion-weighted MRI in diffuse large B-cell lymphoma and follicular lymphoma. *PLoS ONE*, volume 9(1):1–8. ISSN 19326203.

[115] Rakheja, R., Chandarana, H., DeMello, L., et al. (2013). Correlation between standardized uptake value and apparent diffusion coefficient of neoplastic lesions evaluated with whole-body simultaneous hybrid PET/MRI. *American Journal of Roentgenology*, volume 201(5):1115–1119. ISSN 0361803X.

[116] Daisne, J.F., Duprez, T., Weynand, B., et al. (2004). Tumor Volume in Pharyngolaryngeal Squamous Cell Carcinoma: Comparison at CT, MR Imaging, and FDG PET and Validation with Surgical Specimen. *Radiology*, volume 233(1):93–100. ISSN 0033-8419.

[117] De Cobelli, F., Giganti, F., Orsenigo, E., et al. (2013). Apparent diffusion coefficient modifications in assessing gastro-oesophageal cancer response to neoadjuvant treatment: Comparison with tumour regression grade at histology. *European Radiology*, volume 23(8):2165–2174. ISSN 09387994.

[118] Tan, T.H., Boey, C.Y., and Lee, B.N. (2016). Role of Pre-therapeutic (18)F-FDG PET/CT in Guiding the Treatment Strategy and Predicting Prognosis in Patients with Esophageal Carcinoma. *Asia Oceania journal of nuclear medicine & biology*, volume 4(2):59–65. ISSN 2322-5718.

[119] Al-Taan, O.S., Eltweri, A., Sharpe, D., et al. (2014). Prognostic value of baseline FDG uptake on PET-CT in esophageal carcinoma. *World journal of gastrointestinal oncology*, volume 6(5):139–44. ISSN 1948-5204.

[120] Westerterp, M., van Westreenen, H.L., Reitsma, J.B., et al. (2005). Esophageal cancer: CT, endoscopic US, and FDG PET for assessment of response to neoadjuvant therapy–systematic review. *Radiology*, volume 236(3):841–851. ISSN 0033-8419.

[121] Koh, D.m. and Collins, D.J. (2007). Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. *American Journal of Roentgenology*, volume 188(6):1622–1635. ISSN 0361-803X.

[122] Zahra, M.A., Hollingsworth, K.G., Sala, E., Lomas, D.J., and Tan, L.T. (2007). Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy. *The Lancet Oncology*, volume 8(1):63–74. ISSN 14702045.

[123] Heethuis, S.E., van Rossum, P.S., Lips, I.M., et al. (2016). Dynamic contrast-enhanced MRI for treatment response assessment in patients with oesophageal cancer receiving neoadjuvant chemoradiotherapy. *Radiotherapy and Oncology*, volume 120(1):128–135. ISSN 01678140.

[124] Jezzard, P. and Balaban, R.S. (1995). Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic Resonance in Medicine*, volume 34(1):65–73. ISSN 07403194.

[125] Fram, E.K., Herfkens, R.J., Johnson, G.A., et al. (1987). Rapid calculation of T1 using variable flip angle gradient refocused imaging. *Magnetic Resonance Imaging*, volume 5(3):201–208. ISSN 0730725X.

[126] Bol, G.H., Kotte, A.N., van der Heide, U.A., and Lagendijk, J.J. (2009). Simultaneous multi-modality ROI delineation in clinical practice. *Computer Methods and Programs in Biomedicine*, volume 96(2):133–140. ISSN 01692607.

[127] Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, volume 19(6):716–723. ISSN 15582523.

[128] Lei, J., Tian, Y., Zhu, S., et al. (2015). Preliminary study of IVIM-DWI and DCE-MRI in early diagnosis of esophageal cancer. *European Review for Medical and Pharmacological Sciences*, (19):3345–3350.

[129] Weiss, E., Ford, J.C., Olsen, K.M., et al. (2016). Apparent diffusion coefficient (ADC) change on repeated diffusion-weighted magnetic resonance imaging during radiochemotherapy for non-small cell lung cancer: A pilot study. *Lung Cancer*, volume 96:113–119. ISSN 18728332.

[130] Habr-Gama, A., Perez, R.O., Nadalin, W., et al. (2004). Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Annals of surgery*, volume 240(4):711–7; discussion 717–8. ISSN 0003-4932.

[131] Maas, M., Beets-Tan, R.G.H., Lambregts, D.M.J., et al. (2011). Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *Journal of Clinical Oncology*, volume 29(35):4633–4640. ISSN 0732183X.

[132] El Naqa, I., Bradley, J.D., Lindsay, P.E., Hope, A.J., and Deasy, J.O. (2009). Predicting radiotherapy outcomes using statistical learning techniques. *Physics in medicine and biology*, volume 54(18):S9–S30. ISSN 0031-9155.

[133] Oermann, E.K., Rubinsteyn, A., Ding, D., et al. (2016). Using a Machine Learning Approach to Predict Outcomes after Radiosurgery for Cerebral Arteriovenous Malformations. *Scientific reports*, volume 6(February):21161. ISSN 2045-2322 (Electronic).

[134] van Rossum, P.S., Xu, C., Fried, D.V., et al. (2016). The emerging field of radiomics in esophageal cancer: current evidence and future potential. *Translational Cancer Research*, volume 5(4):410–423. ISSN 2219-6803.

[135] Shapiro, J., Van Hagen, P., Lingsma, H.F., et al. (2014). Prolonged time to surgery after neoadjuvant chemoradiotherapy increases histopathological response without affecting survival in patients with esophageal or junctional cancer. *Annals of Surgery*, volume 260(5):807–814. ISSN 15281140.

[136] van der Werf, L.R., Dikken, J.L., van der Willik, E.M., et al. (2018). Time interval between neoadjuvant chemoradiotherapy and surgery for oesophageal or junctional cancer: A nationwide study. *European Journal of Cancer*, volume 91:76–85. ISSN 18790852.

[137] Qiu, B., Wang, D., Yang, H., et al. (2016). Combined modalities of magnetic resonance imaging, endoscopy and computed tomography in the evaluation of tumor responses to definitive chemoradiotherapy in esophageal squamous cell carcinoma. *Radiotherapy and Oncology*, volume 121(2):239–245. ISSN 1879-0887.

[138] Jia, Q.J., Zhang, S.X., Chen, W.B., et al. (2014). Initial experience of correlating parameters of intravoxel incoherent motion and dynamic contrast-enhanced magnetic resonance imaging at 3.0 T in nasopharyngeal carcinoma. *European Radiology*, pp. 3076–3087. ISSN 09387994.

[139] Dyvorne, H., Jajamovich, G., Kakite, S., Kuehn, B., and Taouli, B. (2014). Intravoxel incoherent motion diffusion imaging of the liver : Optimal b-value subsampling and impact on parameter precision and reproducibility. *European Journal of Radiology*, volume 83:2109–2113.

[140] Wurnig, M.C., Germann, M., and Boss, A. (2017). Is there evidence for more than two diffusion components in abdominal organs? - A magnetic resonance imaging study in healthy volunteers. *NMR in Biomedicine*, (September):e3852. ISSN 09523480.

[141] van der Bel, R., Gurney-Champion, O.J., Froeling, M., et al. (2017). A tri-exponential model for intravoxel incoherent motion analysis of the human kidney: In silico and during pharmacological renal perfusion modulation. *European Journal of Radiology*, volume 91(March):168–174. ISSN 18727727.

[142] Guyader, J.M., Bernardin, L., Douglas, N.H., et al. (2014). Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen. *Journal of Magnetic Resonance Imaging*, volume 42:315:330. ISSN 10531807.

[143] Lee, E.Y.P., Hui, E.S.K., Chan, K.K.L., et al. (2015). Relationship between intravoxel incoherent motion diffusion-weighted MRI and dynamic contrast-enhanced MRI in tissue perfusion of cervical cancers. *Journal of Magnetic Resonance Imaging*, volume 42(2):454–459. ISSN 10531807.

[144] Luciani, A., Vignaud, A., and Cavet, M. (2008). Liver Cirrhosis : Intravoxel Incoherent Motion MR Imaging Pilot Purpose : Methods : Results : Conclusion. *Radiology*, volume 249(3):891–899. ISSN 1527-1315.

[145] Zheng, H., Ren, W., Pan, X., et al. (2018). Role of intravoxel incoherent motion MRI in early assessment of the response of esophageal squamous cell carcinoma to chemoradiotherapy: A pilot study. *Journal of Magnetic Resonance Imaging*, pp. 1–10. ISSN 10531807.

[146] Huang, Y.C., Chen, T.w., Zhang, X.M., et al. (2018). Intravoxel incoherent motion diffusion-weighted imaging of resectable oesophageal squamous cell carcinoma: association with tumour stage. *The British Journal of Radiology*, volume 91(1084):20170421. ISSN 0007-1285.

[147] Klein, S., Staring, M., Murphy, K., Viergever, M., and Pluim, J. (2010). elastix: A Toolbox for Intensity-Based Medical Image Registration. volume 29(1):196–205. ISSN 0278-0062.

[148] Lu, W., Jing, H., Ju-Mei, Z., et al. (2017). Intravoxel incoherent motion diffusion-weighted imaging for discriminating the pathological response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Scientific Reports*, volume 7(1):1–9. ISSN 20452322.

[149] Zhu, L., Wang, H., Zhu, L., et al. (2017). Predictive and prognostic value of intravoxel incoherent motion (IVIM) MR imaging in patients with advanced cervical cancers undergoing concurrent chemo-radiotherapy. *Scientific Reports*, volume 7(1):1–9. ISSN 20452322.

[150] Gudbjartsson, H. and Patz, S. (1995). The Rician Distribution of Noisy MRI Data Hákon. *Magnetic Resonance in Medicine*, volume 34(6):910–914. ISSN 15378276.

[151] Heusch, P., Wittsack, H.J., Pentang, G., et al. (2013). Biexponential analysis of diffusion-weighted imaging: comparison of three different calculation methods in transplanted kidneys. *Acta Radiologica*, volume 54(10):1210–1217. ISSN 0284-1851.

[152] Dyvorne, H.a., Galea, N., Nevers, T., et al. (2013). Diffusion-weighted imaging of the liver with multiple b values: effect of diffusion gradient polarity and breathing acquisition on image quality and intravoxel incoherent motion parameters–a pilot study. *Radiology*, volume 266(3):920–9. ISSN 1527-1315.

[153] Jerome, N.P., Orton, M.R., D'Arcy, J.A., et al. (2014). Comparison of free-breathing with navigator-controlled acquisition regimes in abdominal diffusion-weighted magnetic resonance images: Effect on ADC and IVIM statistics. *Journal of Magnetic Resonance Imaging*, volume 39(1):235–240. ISSN 10531807.

[154] Kwee, T.C., Takahara, T., Koh, D.M., Nievelstein, R.A., and Luijten, P.R. (2008). Comparison and reproducibility of ADC measurements in breathhold, respiratory triggered, and free-breathing diffusion-weighted MR imaging of the liver. *Journal of Magnetic Resonance Imaging*, volume 28(5):1141–1148. ISSN 10531807.

[155] Yamashita, H., Kida, S., Sakumi, A., et al. (2011). Four-Dimensional Measurement of the Displacement of Internal Fiducial Markers During 320-Multislice Computed Tomography Scanning of Thoracic Esophageal Cancer. *International Journal of Radiation Oncology Biology Physics*, volume 79(2):588–595. ISSN 03603016.

[156] Dieleman, E.M.T., Senan, S., Vincent, A., et al. (2007). Four-dimensional computed tomographic analysis of esophageal mobility during normal respiration. *International Journal of Radiation Oncology Biology Physics*, volume 67(3):775–780. ISSN 03603016.

[157] Yaremko, B.P., Guerrero, T.M., McAleer, M.F., et al. (2008). Determination of Respiratory Motion for Distal Esophagus Cancer Using Four-Dimensional Computed Tomography. *International Journal of Radiation Oncology Biology Physics*, volume 70(1):145–153. ISSN 03603016.

[158] Feng, M., Balter, J.M., Normolle, D., et al. (2009). Characterization of Pancreatic Tumor Motion Using Cine MRI: Surrogates for Tumor Position Should Be Used With Caution. *International Journal of Radiation Oncology Biology Physics*, volume 74(3):884–891. ISSN 03603016.

[159] Park, S., Farah, R., Shea, S.M., et al. (2017). Evaluation of lung tumor motion management in radiation therapy with dynamic MRI. volume 10135, p. 101351R.

[160] Shi, X., Diwanji, T., Mooney, K.E., et al. (2014). Evaluation of template matching for tumor motion management with cine-MR images in lung cancer patients. *Medical physics*, volume 41(5):052304. ISSN 0094-2405.

[161] Zachiu, C., Denis de Senneville, B., Moonen, C., and Ries, M. (2015). A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: Proof of concept. *Medical physics*, volume 42(7):4137–4148. ISSN 2473-4209 (Electronic).

[162] Roujol, S., Ries, M., Moonen, C., and de Senneville, B.D. (2011). Automatic Nonrigid Calibration of Image Registration for Real Time MR-Guided HIFU Ablations of Mobile Organs. *IEEE Transactions on Medical Imaging*, volume 30(10):1737–1745. ISSN 0278-0062.

[163] Trofimov, A., Vrancic, C., Chan, T.C.Y., Sharp, G.C., and Bortfeld, T. (2008). Tumor trailing strategy for intensity-modulated radiation therapy of moving targets. *Medical Physics*, volume 35(5):1718–1733. ISSN 00942405.

[164] Tang, S., Deville, C., McDonough, J., et al. (2013). Effect of intrafraction prostate motion on proton pencil beam scanning delivery: A quantitative assessment. *International Journal of Radiation Oncology Biology Physics*, volume 87(2):375–382. ISSN 03603016.

[165] Chen, X., Lu, H., Tai, A., et al. (2014). Determination of Internal Target Volume for Radiation Treatment Planning of Esophageal Cancer by Using 4-Dimensional Computed Tomography (4DCT). *International Journal of Radiation Oncology Biology Physics*, volume 90(1):102–109. ISSN 03603016.

[166] van Herk, M., Remeijer, P., Rasch, C., and Lebesque, J.V. (2000). The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *International Journal of Radiation Oncology Biology Physics*, volume 47(4):1121–1135. ISSN 03603016.

[167] Jin, P., Hulshof, M.C., van Wieringen, N., Bel, A., and Alderliesten, T. (2017). Interfractional variability of respiration-induced esophageal tumor motion quantified using fiducial markers and four-dimensional cone-beam computed tomography. *Radiotherapy and Oncology*, volume 124(1):147–154. ISSN 01678140.

[168] Ragnarsdóttir, M. and Kristinsdóttir, E.K. (2006). Breathing Movements and Breathing Patterns among Healthy Men and Women 2069 Years of Age. *Respiration*, volume 73(1):48–54. ISSN 0025-7931.

[169] Parreira, V.F., Bueno, C.J., França, D.C., et al. (2010). Padrão respiratório e movimento toracoabdominal em indivíduos saudáveis: influência da idade e do sexo. *Revista Brasileira de Fisioterapia*, volume 14(5):411–416. ISSN 1413-3555.

[170] Noordman, B.J., Wijnhoven, B.P.L., Lagarde, S.M., et al. (2018). Neoadjuvant chemoradiotherapy plus surgery versus active surveillance for oesophageal cancer: a stepped-wedge cluster randomised trial. *BMC Cancer*, volume 18(1):142. ISSN 1471-2407.

[171] Tofts, P. (1997). Modeling tracer kinetics in dynamic GdDTPA MR imaging. *Journal of Magnetic Resonance Imaging*, pp. 91–101.

[172] Ye, Z.M., Dai, S.J., Yan, F.Q., et al. (2018). DCE-MRI-Derived Volume Transfer Constant (K trans) and DWI Apparent Diffusion Coefficient as Predictive Markers of Short- and Long-Term Efficacy of Chemoradiotherapy in Patients With Esophageal Cancer. *Technology in Cancer Research & Treatment*, volume 17(38):153303461876525. ISSN 1533-0346.

[173] Klaassen, R., Gurney-Champion, O.J., Wilmink, J.W., et al. (2018). Repeatability and correlations of dynamic contrast enhanced and T2* MRI in patients with advanced pancreatic ductal adenocarcinoma. *Magnetic Resonance Imaging*, volume 50(February):1–9. ISSN 0730725X.

[174] Galbán, C.J., Hoff, B.A., Chenevert, T.L., and Ross, B.D. (2017). Diffusion MRI in early cancer therapeutic response assessment. *NMR in Biomedicine*, volume 30(3):e3458. ISSN 09523480.

[175] Mahajan, A., Deshpande, S.S., and Thakur, M.H. (2017). Diffusion magnetic resonance imaging: A molecular imaging tool caught between hope, hype and the real world of personalized oncology. *World Journal of Radiology*, volume 9(6):253. ISSN 1949-8470.

[176] Li, Q.W., Qiu, B., Wang, B., et al. (2018). Prediction of pathologic responders to neoadjuvant chemoradiotherapy by diffusion-weighted magnetic resonance imaging in locally advanced esophageal squamous cell carcinoma: a prospective study. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus*, volume 31(2):1–7. ISSN 1442-2050.

[177] Giganti, F., Ambrosi, A., Petrone, M.C., et al. (2016). Prospective comparison of MR with diffusion-weighted imaging, endoscopic ultrasound, MDCT and positron emission tomography-CT in the pre-operative staging of oesophageal cancer: results from a pilot study. *The British Journal of Radiology*, volume 89(1068):20160087. ISSN 0007-1285.

[178] Usuda, K., Funasaki, A., Sekimura, A., et al. (2018). FDG-PET/CT and diffusion-weighted imaging for resected lung cancer: correlation of maximum standardized uptake value and apparent diffusion coefficient value with prognostic factors. *Medical Oncology*, volume 35(5):66. ISSN 1357-0560.

[179] Bernstine, H., Domachevsky, L., Nidam, M., et al. (2018). 18F-FDG PET/MR imaging of lymphoma nodal target lesions. *Medicine*, volume 97(16):e0490. ISSN 0025-7974.

[180] Virostko, J., Hainline, A., Kang, H., et al. (2017). Dynamic contrast-enhanced magnetic resonance imaging and diffusion-weighted magnetic resonance imaging for predicting the response of locally advanced breast cancer to neoadjuvant therapy: a meta-analysis. *Journal of Medical Imaging*, volume 5(01):1. ISSN 2329-4302.

[181] Goense, L., Borggreve, A.S., Heethuis, S.E., et al. (2018). Patient perspectives on repeated MRI and PET/CT examinations during neoadjuvant treatment of oesophageal cancer. *The British Journal of Radiology*, (February):20170710. ISSN 0007-1285.

[182] Lai, V., Ho Fun Lee, V., On Lam, K., et al. (2017). Intravoxel incoherent motion MR imaging in nasopharyngeal carcinoma: comparison and correlation with dynamic contrast enhanced MR imaging. *Oncotarget*, volume 8(40):68472–68482. ISSN 1949-2553.

[183] Yuan, M., Zhang, Y.D.D., Zhu, C., et al. (2015). Comparison of intravoxel incoherent motion diffusion-weighted MR imaging with dynamic contrast-enhanced MRI for differentiating lung cancer from benign solitary pulmonary lesions. *J Magn Reson Imaging*, volume 43:669–679. ISSN 10531807.

[184] Jiang, L., Lu, X., Hua, B., et al. (2017). Intravoxel Incoherent Motion Diffusion-Weighted Imaging Versus Dynamic Contrast-Enhanced Magnetic Resonance Imaging: Comparison of the Diagnostic Performance of Perfusion-Related Parameters in Breast. *Journal of computer assisted tomography*, volume 00(00):1–6. ISSN 1532-3145 (Electronic).

[185] Li, F.P., Wang, H., Hou, J., et al. (2018). Utility of intravoxel incoherent motion diffusion-weighted imaging in predicting early response to concurrent chemoradiotherapy in oesophageal squamous cell carcinoma.

[186] Skórska, M. and Piotrowski, T. (2017). Personalized radiotherapy treatment planning based on functional imaging. *Reports of Practical Oncology & Radiotherapy*, volume 22(4):327–330. ISSN 15071367.

[187] Abedi, I., Tavakkoli, M.B., Jabbari, K., Amouheidari, A., and Yadegarfard, G. (2017). Dosimetric and Radiobiological Evaluation of Multiparametric MRI-Guided Dose Painting in Radiotherapy of Prostate Cancer. *Journal of medical signals and sensors*, volume 7(2):114–121. ISSN 2228-7477.

[188] Grönlund, E., Johansson, S., Nyholm, T., Thellenberg, C., and Ahnesjö, A. (2018). Dose painting of prostate cancer based on Gleason score correlations with apparent diffusion coefficients. *Acta Oncologica*, volume 57(5):574–581. ISSN 0284-186X.

[189] Liu, F., Xi, X.p., Wang, H., et al. (2017). PET/CT-guided dose-painting versus CT-based intensity modulated radiation therapy in locoregional advanced nasopharyngeal carcinoma. *Radiation Oncology*, volume 12(1):15. ISSN 1748-717X.

[190] Ye, Z., Fang, J., Dai, S., et al. (2017). Inter- and intra-observer reproducibility of ADC measurements in esophageal carcinoma primary tumors. *Oncotarget*, volume 8(54):92880–92889. ISSN 1949-2553.

[191] Sorace, A.G., Wu, C., Barnes, S.L., et al. (2018). Repeatability, reproducibility, and accuracy of quantitative mri of the breast in the community radiology setting. *Journal of Magnetic Resonance Imaging*, pp. 1–13. ISSN 10531807.

[192] Bobek, V., Matkowski, R., Gürlich, R., et al. (2014). Cultivation of circulating tumor cells in esophageal cancer. *Folia Histochemica et Cytobiologica*, volume 52(3):171–177. ISSN 1897-5631.

[193] Song, Y., Tian, T., Shi, Y., et al. (2017). Enrichment and single-cell analysis of circulating tumor cells. *Chemical Science*, volume 8(3):1736–1751. ISSN 2041-6520.

[194] Hoeppner, J. and Kulemann, B. (2017). Circulating tumor cells in esophageal cancer. *Oncology Research and Treatment*, volume 40(7-8):417–422. ISSN 22965262.

[195] Qiao, Y., Li, J., Shi, C., et al. (2017). Prognostic value of circulating tumor cells in the peripheral blood of patients with esophageal squamous cell carcinoma. *Onco Targets Ther.*, volume 10:1363–1373. ISSN 11786930.

[196] Gründner, J., Prokosch, H.u., Stürzl, M., et al. (2018). Predicting Clinical Outcomes in Colorectal Cancer Using Machine Learning. *Studies in health technology and informatics*, volume 247:101–105. ISSN 0926-9630.

[197] Sun, Y., Reynolds, H.M., Wraith, D., et al. (2018). Voxel-wise prostate cell density prediction using multiparametric magnetic resonance imaging and machine learning. *Acta Oncologica*, volume 0(0):1–7. ISSN 0284-186X.

[198] Ching, T., Himmelstein, D.S., Beaulieu-Jones, B.K., et al. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, volume 15(141):20170387. ISSN 1742-5689.

[199] Hou, Z., Li, S., Ren, W., et al. (2018). Radiomic analysis in T2W and SPAIR T2W MRI: predict treatment response to chemoradiotherapy in esophageal squamous cell carcinoma. *Journal of Thoracic Disease*, volume 10(4):2256–2267. ISSN 20721439.

Addendum

Nederlandse samenvatting

Dankwoord

List of publications

Curriculum vitae auctoris

Nederlandse samenvatting

Wereldwijd staat slokdarmkanker op de 8ste plek van meest voorkomende kankersoorten en de incidentie blijft stijgen. Slokdarmkanker wordt gekenmerkt door zeer slechte overlevingskansen, met de vijfjaars totale overleving tussen de 15 en 25%. Door het invoeren van een totale slokdarmresectie als behandeling zijn deze overlevingskansen gestegen tot ongeveer 35%. Tegenwoordig behoort het geven van een neoadjuvante therapie (wat een aantal weken voorafgaand aan de operatie wordt gegeven) tot de standaardbehandeling waardoor de overlevingskansen met nog eens 13% gestegen zijn. De neoadjuvante therapie duurt vijf weken en bestaat uit wekelijkse toediening van chemotherapie gecombineerd met een radiotherapieschema verdeeld over 23 fracties. Door na de operatie het resectiepreparaat te onderzoeken is inmiddels bekend dat in 29% van de patiënten geen tumorrest wordt gevonden en dat dus een complete respons is behaald. Voor deze patiënten zijn de overlevingskansen een stuk positiever met vijfjaars overlevingskansen variërend van 48 tot 65%. Neoadjuvante therapie is echter een zware behandeling en ook de slokdarmresectie is een zeer ingrijpende operatie, die substantiële risico's op complicaties met zich meebrengt. Aangezien het bekend is dat niet alle patiënten in gelijke mate profiteren van de neoadjuvante therapie zouden we graag in staat zijn te voorspellen welke patiënten een mogelijke complete respons gaan bereiken en bij welke patiënten de neoadjuvante therapie niet aan lijkt te slaan. In het geval van een verwachte complete respons zouden we in de toekomst graag de operatie willen voorkomen. Bij patiënten waar echter geen respons waargenomen wordt, kan mogelijk de neoadjuvante therapie aangepast of zelfs onderbroken worden, waarna bijvoorbeeld de operatie sneller plaats zou kunnen vinden.

Met behulp van *magnetic resonance imaging* (MRI) kunnen verschillende soorten scans gemaakt worden, waarvan de anatomische scans het meest bekend zijn. Echter, met MRI kan bijvoorbeeld ook de perfusie of diffusie van het weefsel gevisualiseerd worden, in welk geval we praten over functionele MRI-scans. Naar het gebruik van functionele MRI voor responspredictie wordt veel onderzoek gedaan. Voor slokdarmtumoren is in een aantal (kleine) studies al aangetoond dat functionele MRI een bijdrage kan leveren aan het voorspellen van respons op neoadjuvante therapie.

Om in de toekomst voor elke patiënt een behandelstrategie 'op maat' te kunnen maken kan op verschillende vlakken verbetering behaald worden. Ten eerste is het van belang een betrouwbare methode te hebben om de respons van patiënten met slokdarmkanker te voorspellen. In dit proefschrift ligt de focus voornamelijk op het onderzoeken van verschillende functionele MRI-scans en hun mogelijke waarde in het voorspellen van de respons bij patiënten met slokdarmkanker. Ook wordt onderzocht hoe het tegelijkertijd inzetten van verschillende MRI-scans van toegevoegde waarde kan zijn. Daarnaast onderzoeken we de mogelijke toegevoegde waarde van MRI met een klinisch veel gebuikte PET/CT-scan. Ten tweede, als patiënten in de toekomst behandeld worden

met de nieuwe MRI-geleide radiotherapie (MR-Linac) dan kan tijdens de bestraling de tumor 'gevolgd' worden waardoor de tumor met hogere nauwkeurigheid bestraald kan worden. Hierdoor verwachten we dat de op populatie-gebaseerde behandelingsmarges gereduceerd kunnen worden. De mogelijke impact van deze vernieuwde behandeling is nog onbekend. Als eerste stap is in dit proefschrift onderzocht wat de beweging is van slokdarmtumoren op verschillende tijdsschalen.

In **Hoofdstuk 2** is onderzocht of *dynamic contrast-enhanced* (DCE) MRI gebruikt kan worden om de respons bij patiënten met slokdarmkanker te voorspellen. Bij deze techniek worden na injectie van een contrastmiddel in de bloedbaan herhaaldelijk scans gemaakt, waardoor het gedrag van het contrastmiddel in de slokdarmtumor gemeten kan worden. Voor een kwantitatieve vergelijking van waardes tussen scans die op verschillende tijdstippen genomen zijn, kan de daadwerkelijke concentratie aan contrast in het weefsel berekend worden. In deze studie vonden we dat veranderingen over tijd in de geaccumuleerde concentratie in de tumor, gemeten gedurende 60 seconden vanaf het moment van instroming van het contrastmiddel, een voorspellende maat bleek te zijn voor respons. Er werd derhalve geconcludeerd dat DCE-MRI van toegevoegde waarde kan zijn in een respons-predictiemodel.

In een eerdere studie werd al gevonden dat *diffusion-weighted* (DW) MRI respons kan voorspellen bij patiënten met slokdarmkanker. Bij DW-MRI is het signaal op de scan afhankelijk van de bewegelijkheid van watermoleculen in het weefsel. In een scan met hoge diffusieweging, wordt er een hoog signaal gevonden in de scan indien er weinig beweging tussen de watermoleculen mogelijk is. Dit is het geval bij weefsel met een hoge celdichtheid, zoals bijvoorbeeld tumorweefsel. Als een behandeling aanslaat wordt verwacht dat de celdichtheid zich begint te herstellen en dat dus het signaal lager wordt. Voor een kwantitatieve beoordeling van het diffusie-gewogen MRI-signaal wordt er gekeken naar het verschil in signaal tussen een scan met een lage en een scan met een hoge diffusieweging. Hieruit wordt de diffusiecoëfficiënt ADC berekend, welke omgekeerd evenredig is aan het signaal (tumorweefsel vertoont dus een lage ADC waarde). In **Hoofdstuk 3** is onderzocht wat de mogelijke toegevoegde waarde is van de ADC-parameter zoals gemeten in DW-MRI en de SUV-parameter zoals gemeten in *^{18}F-fluorodeoxyglucose positron emission tomography* met geïntegreerde *computed tomography* (^{18}F-FDG-PET/CT). SUV staat voor *standardized uptake value* en is een maat voor de concentratie radioactiviteit in het weefsel. Bij deze techniek wordt via de bloedbaan radioactief-gelabelde glucose toegediend. Door de hoge stofwisseling in tumoren wordt hier veel glucose verbruikt waardoor tumoren als een hotspot zichtbaar zijn op een PET scan. In deze studie werden geen correlaties gevonden tussen ADC en SUV. Dit betekent dat beide technieken van toegevoegde waarde zouden kunnen zijn in een model waarin respons voorspeld wordt en dat het gebruik van beide technieken in toekomstige studies verantwoord is.

In **Hoofdstuk 4** is de toegevoegde waarde van DCE-MRI met DW-MRI onder-

zocht. Aangezien beide technieken andere tumor karakteristieken visualiseren (perfusie en diffusie) zou het combineren van deze technieken mogelijk van toegevoegde waarde kunnen zijn. In deze studie vonden we dat beide technieken inderdaad van toegevoegde waarde zijn en dat een combinatie hogere voorspellende waardes geeft, in vergelijking met het gebruik van slechts één techniek. Er werd echter in deze studie wel gevonden dat de DW-MRI-scans niet van voldoende kwaliteit waren omdat er veel bewegingsartifacten waargenomen werden in de scans. Daarom is na deze studie het scanprotocol aangepast. Bij deze wijziging hebben we ook het aantal diffusiewegingen aangepast. In onze eerdere studies werd de ADC bepaald met behulp van 3 diffussiewaardes, waarbij in het nieuwe protocol 13 waardes zijn opgenomen. Het is bekend dat het signaalverloop voor lage diffusiewaardes gedomineerd wordt door perfusie, waardoor het signaalverloop gesplitst kan worden in een snelle perfusie-gedomineerde diffusie en een langzame diffusie. Het toevoegen van meer diffusiewaardes stelt ons daarom in staat om uitgebreidere analyses op het signaal te kunnen uitvoeren. Deze techniek is ook wel bekend als *intravoxel incoherent motion* (IVIM) MRI. In **Hoofdstuk 5** hebben we verschillende IVIM fitting-methodes vergeleken voor een grote dataset met slokdarmtumoren. In deze studie vonden we dat ondanks dat de determinatiecoëfficiënt R^2, een maat voor de 'goedheid' van de fit, vergelijkbaar was voor de verschillende methodes, de gefitte parameters wel degelijk verschillen lieten zien. Geconcludeerd werd dat bij het rapporteren van IVIM-MRI resultaten het van belang is dat uitgebreide informatie over de toegepaste fittingprocedures bijgevoegd wordt.

Zoals eerder genoemd kan er mogelijk veel winst behaald worden in het personaliseren van de behandeling als MRI-gestuurde radiotherapie gebruikt gaat worden. Naast het feit dat marges verkleind zouden kunnen worden, omdat de tumor gevolgd kan worden tijdens de bestraling, maakt het ook *dose-painting* mogelijk. Hierbij kunnen binnen een tumor delen van de tumor met hogere dosis bestraald worden, als deze delen er bijvoorbeeld van verdacht worden niet voldoende te reageren op de behandeling. In **Hoofdstuk 6** is de intra-fractie beweging van slokdarmtumoren onderzocht in verschillende tijdframes. Zo werd er gedurende de 5 weken van neoadjuvante therapie wekelijks een MRI-scan gemaakt. Binnen deze MRI-scan bevonden zich twee series met cine-MRI-scans. Een cine-MRI-scan is een 2D-scan die op hoge tijdsresolutie opgenomen wordt, met in onze studie 75 frames binnen 45 seconden. Deze scan werd herhaald na 10 minuten. Hierdoor kon zowel binnen 45 seconden de beweging geanalyseerd worden, als de mogelijke drift van de tumor in 10 minuten. Zoals ook in andere studies gevonden is werd in de cranio-caudale richting de grootste beweging gevonden, veroorzaakt door de ademhaling, met een gemiddelde maximale amplitude van 12.7 mm. Uit verder onderzoek van deze beweging bleek dat er erg grote verschillen gevonden werden tussen de bewegingsamplitudes zowel binnen patiënten als tussen patiënten door de tijd heen. Verder werd er gevonden dat er met enige regelmaat een tumordrift kan ontstaan wat grote gevolgen zou kunnen hebben indien dit ook tijdens een bestraling plaatsvindt. In deze studie werd geconcludeerd dat identificatie van de

beweging voor de start van een behandeling niet representatief is voor de daadwerkelijke beweging tijdens een behandeling en dat real-time volgen van de tumor tijdens radiotherapie een vereiste is voor het veilig reduceren van de behandelingsmarges en het eventueel toepassen van techieken als *dose-painting*.

Concluderend is ons uiteindelijke toekomstdoel het ontwikkelen van een gepersonaliseerde behandeling voor patiënten met slokdarmkanker. Hierin willen we dat de marges verkleind worden om toxiciteit in omliggende weefsels te kunnen reduceren en daarnaast *dose-painting* mogelijk maken door het real-time volgen van de tumor tijdens bestraling. Als dit samengaat met een model dat nauwkeurig de respons kan voorspellen, kunnen we voor patiënten die een goede respons bereiken de ingrijpende operatie voorkomen. In dit proefschrift hebben we de potentie laten zien van DCE-MRI voor het voorspellen van respons. Daarnaast hebben we laten zien dat het combineren van meerdere functionele MRI-scans, eventueel gecombineerd met PET/CT, van toegevoegde waarde is. Dit kan gebruikt gaan worden in een multi-parametrisch multi-modaliteiten voorspellingsmodel. Verder hebben we laten zien dat, vanwege het willekeurig gedrag van de variaties in de beweging van slokdarmtumoren, het belangrijk is om het real-time volgen van een tumor mogelijk te maken en om zodoende veilig de marges te kunnen reduceren.

Dankwoord

Tot slot zou ik graag iedereen willen bedanken die op enigerlei wijze de afgelopen vier jaar een bijdrage heeft geleverd aan mijn promotie en het tot stand komen van dit proefschrift.

Allereerst wil ik alle patiënten bedanken die deelnamen aan de verschillende studies waarop dit proefschrift gebasseerd is. In een al heel zware tijd in hun leven kiezen ze ervoor verschillende extra MRI scans te ondergaan om zo hun bijdrage te kunnen leveren aan de behandeling voor toekomstige patiënten met slokdarmkanker, hiervoor heb ik veel bewondering.

Dank aan mijn promotor Jan, voor het bieden van de kans om een promotietraject uit te voeren op jouw afdeling. Ik heb hier ontzettend veel geleerd en met heel veel plezier vier jaar gewerkt. Bedankt voor de kans om meerdere (buitenlandse) congressen bij te wonen tijdens mijn promotie, naast het feit dat dit natuurlijk altijd heel gezellig was is het ook ontzettend leerzaam geweest.

Astrid, ik wil jou in het bijzonder bedanken voor je begeleiding gedurende mijn promotie. Jij hebt mij ontzettend veel kunnnen leren over het doen van onderzoek en bent altijd erg betrokken geweest bij mijn persoonlijke ontwikkeling als onderzoeker. Ook al was je zelf druk met je opleiding, je maakte altijd tijd vrij en bij tegenvallende resultaten wist jij me te motiveren om door te gaan en er 'dieper in te duiken'. Bedankt voor de fijne samenwerking. Ik wens je heel veel succes in je toekomstige baan als klinish fysicus.

Beste leden van de beoordelingscommissie, prof. dr. I.H.M. Borel Rinkes, prof. dr. C.T.W. Moonen, prof. dr. H.W.A.M. de Jong, prof. dr. S. Both, prof. dr. ir. M.J.P. van Osch, hartelijk dank voor het lezen en beoordelen van dit proefschrift en voor het plaatsnemen in de oppositie.

Ook wil ik Marielle en Stella bedanken voor hun begeleiding tijdens Astrid haar respectievelijk 1ste en 2de zwangerschapsverlof. Jullie hebben beiden een erg drukke agenda en wisten toch tijd vrij te maken om resultaten te bespreken en te kijken welke kant we daarmee verder op konden.

Ik wil graag de hele oes-groep bedanken voor de fijne samenwerking. Astrid, Gert, Stella, Onne, Irene, Peter, Alicia, Lucas, Mick en natuurlijk ook Richard en Jelle, bedankt voor jullie bijdrage en input in mijn onderzoek de afgelopen jaren. Lucas en Alicia in het bijzonder voor de fijne samenwerking, we waren een mooie aanvulling op elkaar en we hebben er met al die iteraties intekeningen aardig wat gezamenlijke uurtjes opzitten.

De afdeling kent een ontzettend grote groep OiO's. Bedankt voor alle leuke uitjes en congressen samen en voor jullie feedback op presentaties bij een van de gezamenlijke

group-meetings. Ook de 7T groep wil ik bedanken, Nikki, Quincy, Arjan en Jolanda, op de werkvloer kwamen we elkaar nauwelijks tegen maar op de congressen hadden we het altijd erg gezellig.

Dank aan alle MRI-laboranten van de radiotherapie voor het scannen van alle patiënten. In het bijzonder wil ik graag Tuan en Ellart bedanken voor de fijne samenwerking afgelopen jaren en het meedenken bij wijzigingen in het scanprotocol.

Bedankt aan alle co-auteurs voor hun bijdrage aan de manuscripten. Bedankt voor de samenwerking met het AVL/NKI in Amsterdam en het MD Anderson in Houston, Texas, USA. In de toekomst wordt deze samenwerking uitgebreid met nog een aantal andere ziekenhuizen. Ik ben heel benieuwd wat voor moois er uit de toekomstige studie(s) gaat komen.

Aan alle kamergenootjes die ik door de tijd heen heb gehad, bedankt. Een fijne werkomgeving hebben is een belangrijk onderdeel van het met plezier naar werk gaan, dankzij de goede sfeer in de kamer en gezamenlijke lunches is dit zeker het geval geweest. In het bijzonder wil ik Joris en Hans bedanken. We hebben (bijna) vanaf het begin een kamer gedeeld, en wel de mooiste van het ziekenhuis! We hadden snel een klik en omdat we in dezelfde fase van onze promoties zaten begrepen we elkaars pieken en dalen extra goed. Bedankt voor het bijspringen als 'bad-eendje', voor het meedenken bij mijn onderzoek, voor het nalezen van abstracts en artikelen, maar bovenal voor vier hele gezellige jaren samen. Hans, wij hebben samen heel wat Matlab en Elastix 'uitdagingen' kunnen tackelen. Joris, naast je medische kennis die regelmatig van pas kwam heb ik ook van je programmeerkennis veel kunnen opsteken. We zijn een maand na elkaar gestart en promoveren ook een maand na elkaar (dan wel in andere volgorde). Het was een veelbewogen traject en ik wil je bedanken dat je mijn paranimf wilt zijn tijdens mijn verdediging.

Ragna, inmiddels al meer dan 12 jaar beste vriendinnen, we kennen elkaar door en door en hebben soms aan weinig woorden genoeg om elkaar te begrijpen. Bedankt voor het zijn van een luisterend oor als dat nodig was, en natuurlijk voor alle momenten samen.

Lieve familie en schoonfamilie, bedankt voor jullie grenzeloze vertrouwen in mijn kunnen, jullie steun als het even tegenviel en het zijn van een luisterend oor. Ik ben jullie allemaal ontzettend dankbaar. Anna, zeker in de laatste maanden toen mijn stresslevels zo nu en dan piekte was jij slechts 'one call away' en hier heb ik regelmatig gebruik van gemaakt. Lieve zus, bedankt voor je steun en lieve woorden.

Lieve Robbie, jij hebt mij de ogen geopend voor de mogelijkheid om een PhD te doen en nu, een kleine vijf jaar later, sta je aan mijn zijde als paranimf. Bedankt voor je steun en positiviteit als ik dat nodig had, je input bij het vinden van oplossingen als ik vast zat, en bovenal natuurlijk bedankt voor alle fijne momenten samen.

List of publications

Published papers

Sophie E. Heethuis, Peter S.N. van Rossum, Irene M. Lips, Lucas Goense, Francine E. Voncken, Onne Reerink, Richard van Hillegersberg, Jelle P. Ruurda, Marielle E. Philippens, Marco van Vulpen, Gert J. Meijer, Jan J.W. Lagendijk, Astrid L.H.M.W. van Lier. (2016) Dynamic contrast-enhanced MRI for treatment response assessment in patients with oesophageal cancer receiving neoadjuvant chemo-radiotherapy. Radiotherapy and Oncology, 120(1), 128-135. doi:10.1016/j.radonc.2016.05.009

Lucas Goense, **Sophie E. Heethuis**, Peter S.N. van Rossum, Francine E. Voncken, Jan J.W. Lagendijk, Marnix G.E.H. Lam, Chris H. Terhaard, Richard van Hillegersberg, Jelle P.Ruurda, Stella Mook, Astrid L.H.M.W. van Lier, Steven H. Lin and Gert J. Meijer (2018) Correlation between functional imaging markers derived from diffusion-weighted MRI and 18F-FDG PET/CT in esophageal cancer. Nuclear Medicine Communications, 39(1):6067
doi:10.1097/MNM.0000000000000771

Sophie E. Heethuis, Lucas Goense, Peter S.N. van Rossum, A.S. Borggreve, S. Mook, Francine E.M. Voncken, Annemarieke Rutten, Berthe M.P. Aleman, Richard van Hillegersberg, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier (2018) DW-MRI and DCE-MRI are of complementary value in predicting pathologic response to neoadjuvant chemoradiotherapy for esophageal cancer. Acta Oncologica
doi:10.1080/0284186X.2018.1473637

Sophie E. Heethuis, A.S. Borggreve, Lucas Goense, Peter S.N. van Rossum, S. Mook, Richard van Hillegersberg, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier (2018) Quantification of variations in intra-fraction motion of esophageal tumors over the course of neoadjuvant chemoradiotherapy based on cine-MRI. Physics in Medicine and Biology, Vol. 63 (14):145019
doi:10.1088/1361-6560/aacfb5

Submitted papers

Sophie E. Heethuis, A.S. Borggreve, Lucas Goense, Peter S.N. van Rossum, S. Mook, Richard van Hillegersberg, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier (2018) Intravoxel incoherent motion (IVIM) diffusion-weighted magnetic resonance imaging of esophageal cancer: analysis of three fitting approaches of a bi-exponential model.

Additional papers

Lucas Goense, Alicia Borggreve, **Sophie Heethuis**, Astrid van Lier, Marnix Lam, Richard van Hillegersberg, Stella Mook, Gert Meijer, Peter van Rossum and Jelle Ruura (2017) Patient perspectives on repeated MRI and PET/CT examinations during neoadjuvant treatment of oesophageal cancer. The British Journal of Radiology (2018) 91:20170710. doi: 10.1259/bjr.20170710

Silvia Wognum*, **Sophie E. Heethuis***, Tezontl Rosario, Mischa S. Hoogeman, Arjan Bel (2014) Validation of deformable image registration algorithms on CT images of ex vivo porcine bladders with fiducial markers. Medical Physics 41:71916. doi: 10.1118/1.4883839 **shared first authorship*

Abstracts and conference proceedings

Sophie E. Heethuis, A.S. Borggreve, Lucas Goense, Peter S.N. van Rossum, S. Mook, Richard van Hillegersberg, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk and Astrid L.H.M.W. van Lier. Quantifying intra-fraction motion in esophageal tumors throughout nCRT based on cine-MRI.
E-poster at ESTRO 37, 2018

Sophie E. Heethuis, Lucas Goense, Peter S.N. van Rossum, Alicia S. Borggreve, Stella Mook, Francine E. Voncken, Richard van Hillegersberg, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk, Astrid L.H.M.W. van Lier. Combining DW- and DCE-MRI for treatment response assessment in patients with esophageal cancer undergoing neoadjuvant chemo-radiotherapy.
Poster at ISMRM, 2017

Sophie E. Heethuis, Lucas Goense, Alicia S. Borggreve, Peter S.N. van Rossum, Richard van Hillegersberg, Jelle P. Ruurda, Stella Mook, Gert J. Meijer, Jan J.W. Lagendijk, Astrid L.H.M.W. van Lier. Can diaphragm motion function as a surrogate for motion of esophageal tumors during treatment?
Poster at ISMRM, 2017 and E-poster at ESTRO 36, 2017

Sophie E. Heethuis, Lucas Goense, Peter S.N. van Rossum, Irene M. Lips, Richard van Hillegersberg, Jelle P. Ruurda, Marco van Vulpen, Gert J. Meijer, Jan J.W. Lagendijk, Astrid L.H.M.W. van Lier. The feasibility of performing intravoxel incoherent motion for esophageal cancer and an initial comparison with dynamic contrast-enhanced MRI.
Poster at ISMRM, 2016 and Oral presentation at ISMRM Benelux, 2016

Sophie E. Heethuis, Peter S.N. van Rossum, Irene M. Lips, Marco van Vulpen, Marielle E.P. Philippens, Jelle P. Ruurda, Gert J. Meijer, Jan J.W. Lagendijk, Astrid L.H.M.W. van Lier. Potential of DCE-MRI for treatment response assessment in esophageal cancer.
Oral at both 3th ESTRO forum, 2015 and ISMRM Benelux, 2015

Sophie E. Heethuis, Peter S.N. van Rossum, Irene M. Lips, Jan J.W. Lagendijk, Astrid L.H.M.W. van Lier. Comparison of Image Registration Strategies to Improve DCE-MRI Uptake Curves in Esophageal Cancer.
Oral presentation at AAPM meeting, 2014

Authors and affiliations

University Medical Center Utrecht, Utrecht, The Netherlands
Department of Radiotherapy
Astrid L.H.M.W. van Lier, PhD
Jan J.W. Lagendijk, PhD
Gert J. Meijer, PhD
Stella Mook, MD, PhD
Irene M. Lips, MD, PhD
Onne Reerink, MD, PhD
Marielle E.P. Philippens, MD, PhD
Marco van Vulpen, MD, PhD
Chris H. Terhaard, MD, PhD

Department of Surgery
Lucas Goense, MD
Peter S.N. van Rossum, MD, PhD
Alicia. S. Borggreve, MD
Richard van Hillegersberg, MD, PhD
Jelle P.Ruurda, MD, PhD

Department of Radiology and Nuclear Medicine
Marnix G.E.H. Lam, MD, PhD

The Netherlands Cancer Institute, Amsterdam, The Netherlands
Department of Radiotherapy
Francine E. Voncken, MD
Annemarieke Bartels-Rutten, MD
Berthe M.P. Aleman, MD, PhD

The University of Texas MD Anderson Cancer Center, Houston (TX), United States of America
Department of Radiation Oncology
Steven H. Lin, MD, PhD

Curriculum vitae auctoris

Sophie Elise Heethuis was born in Alkmaar, The Netherlands, on 20 May 1990. She attended atheneum at Jan Arentsz College in Alkmaar, The Netherlands. In 2008 she started the bachelor's curriculum *Natuur- en Sterrenkunde* at the University of Amsterdam, the Netherlands, which she completed in 2011 with a thesis on *Assessment of pulse wave velocity in coronary arteries*. She continued at the University of Amsterdam to follow the master programme Physics (track Life and Health). Her final master research under supervision of Dr. A. Bel and Dr. Ir. S. Wognum at the Academic Medical Centre (AMC) in Amsterdam, The Netherlands, was titled *Ex vivo validation of non-rigid image registration for bladder structures* and was completed in 2013. In december 2013 Sophie started her PhD at University Medical Center Utrecht, under supervision of Prof. Dr. Ir. J.J.W. Lagendijk and Dr. Ir. A.L.H.M.W. van Lier. Her research, the results of which are described in this thesis, focuses on methods for MR-based response monitoring for patients with esophageal cancer undergoing neoadjuvant chemoradiotherapy.