

UvA-DARE (Digital Academic Repository)

MRI of pancreatic cancer for radiotherapy

Gurney-Champion, O.J.

Publication date

2017

Document Version

Final published version

License

Other

[Link to publication](#)

Citation for published version (APA):

Gurney-Champion, O. J. (2017). *MRI of pancreatic cancer for radiotherapy*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

MRI of pancreatic cancer for radiotherapy



Oliver Jacob Gurney-Champion

MRI of pancreatic cancer for radiotherapy

Oliver Jacob Gurney-Champion

Colophon

ISBN: 978-94-028-0546-8

Printing: Ipskamp printing.

Typesetting: Legatron Electronic Publishing.

Design and lay-out: Tom van Hoogstraten.

Copyright © 2017, O.J. Gurney-Champion, Amsterdam, The Netherlands

The copyright of some of the articles that have been accepted for publication, published or submitted, has been transferred to the respective journals.

MRI of pancreatic cancer for radiotherapy

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op woensdag 29 maart 2017, te 14:00 uur

door

Oliver Jacob Gurney-Champion
geboren te Londen, Verenigd Koninkrijk

Promotiecommissie

Promotores:	Prof. dr. J. Stoker	Universiteit van Amsterdam
	Prof. dr. C.R.N. Rasch	Universiteit van Amsterdam
Copromotores:	Dr. ir. A.J. Nederveen	Universiteit van Amsterdam
	Dr. A. Bel	Universiteit van Amsterdam
Overige leden:	Prof. dr. ir. G.J. Strijkers	Universiteit van Amsterdam
	Prof. dr. L.J.A. Stalpers	Universiteit van Amsterdam
	Prof. dr. B.W. Raaymakers	Universiteit Utrecht
	Prof. dr. U.A. van der Heide	Leids Universitair Medisch Centrum
	Dr. J.H. Runge	Academisch Medisch Centrum

Faculteit der Geneeskunde

Table of contents

Chapter 1	Introduction and outline	1
Chapter 2	Considerable interobserver variation in delineation of pancreatic cancer on 3DCT and 4DCT: a multi-institutional study; <i>Submitted</i>	19
Chapter 3	Addition of MRI improves interobserver variation in CT-based pancreatic tumor delineation; <i>Submitted</i>	43
Chapter 4	Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI; <i>Medical Physics 2015;42(5):2638–2647</i>	65
Chapter 5	Quantitative assessment of biliary stent artifacts on MR images: Potential implications for target delineation in radiotherapy; <i>Medical Physics 2016;43(10):5603–5615</i>	89
Chapter 6	Revisiting the potential of alternating repetition time balanced steady-state free precession imaging in the abdomen at 3 T; <i>Investigative Radiology 2016;51(9):560–568</i>	123
Chapter 7	Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas; <i>Investigative Radiology 2016;51(4):211–220</i>	145
Chapter 8	How best to fit diffusion-weighted magnetic resonance imaging data of pancreatic cancer patients; <i>Submitted</i>	181
Chapter 9	General discussion	207
&	Summary	226
	Samenvatting	230
	PhD Portfolio	234
	List of Publications	236
	Curriculum Vitae	241
	Dankwoord	242



CHAPTER 1

Introduction and outline

Oliver J Gurney-Champion

Introduction

In the following paragraphs I will elaborate on the title of this thesis: “Magnetic resonance imaging of pancreatic cancer for radiotherapy”. First, pancreatic cancer and treatments for pancreatic cancer are discussed. Then, radiotherapy of pancreatic cancer patients is explained in detail. Several shortcomings of image guided radiotherapy might be overcome by magnetic resonance imaging (MRI). Therefore, MRI, and in particular diffusion-weighted imaging (DWI), is explained and put into the perspective of radiotherapy. Finally, the outline of this thesis is discussed, in which MRI techniques are developed, optimized and validated for radiotherapy of pancreatic cancer patients.

Pancreatic cancer

Pancreatic cancer is one of the deadliest of all major cancers with a median survival of 4.4 months after diagnosis [1]. Pancreatic cancer often (78%) occurs in the head of the pancreas (Fig. 1.1 illustrates pancreatic anatomy), but also can occur in the body (11%) or tail (11%) [2]. The estimated number of mortalities as a result of pancreatic cancer worldwide (330,400 per year) is only marginally smaller than the estimated incidence (337,900 per year) [3]. Therefore, pancreatic cancer is currently ranked as the fourth most common cause of cancer-related deaths in more developed countries, whereas it is only the tenth most common cancer. Despite all technical advances in medicine, the survival rate of pancreatic cancer has not largely improved in the past 30 years [4]. Pancreatic cancer is therefore expected to soon become the third most common cause of cancer-related deaths [5].

The only known potentially curative treatment for pancreatic cancer is surgical resection [6]. Only tumors that do not involve the celiac artery or superior mesenteric artery and show no distant metastases are eligible for resection [7, 8]. Therefore, unfortunately, only 15–20% of pancreatic cancer patients are eligible for surgery at the time of diagnosis [9]. Furthermore, even after a successful resection with no remaining tumor at the edge of the resected specimen, patients still have a poor median survival of 23 months [10].

In addition, chemotherapy, radiotherapy and radiochemotherapy (i.e. the combination of radiotherapy and chemotherapy) are commonly applied. These therapies are used in a palliative setting [11, 12] or alongside surgery to improve tumor control. When these therapies are used alongside surgery, it is either in an adjuvant (after surgery) [13-16] or neoadjuvant (before surgery) [17-21] setting. Despite the indication that (neo)adjuvant radiochemotherapy may improve pancreatic cancer

patients life expectancy [15, 21], may delay recurrent disease [16] and may improve local tumor control [19], there are still many challenges in radiotherapy that should be overcome to improve the treatment outcome further.

Radiotherapy

In radiotherapy, ionizing radiation is used to kill cancer cells. The ionizing radiation can damage DNA in cells, which can cause cell death. Pancreatic cancer patients are typically irradiated by an ionizing beam that originates in a linear accelerator (linac, Fig. 1.2). In such a setup, the radiation penetrates the patient to reach the tumor. Therefore, healthy tissue is inevitably irradiated too. DNA of both tumor and healthy tissue can be damaged from the ionizing radiation. However, healthy tissue repairs the DNA damage more rapidly than tumor tissue. Therefore, radiotherapy is usually administered in multiple daily fractions, whereby the healthy tissue recovers more than the tumorous tissue between each session. Due to the toxicity of irradiation, the goal of radiotherapy is to maximize the dose to the tumor while minimizing the dose to healthy tissue, in particular, radiation sensitive organs at risk (OARs). For pancreatic cancer patients, the OARs include the small bowel, stomach, liver, kidneys and spinal cord.

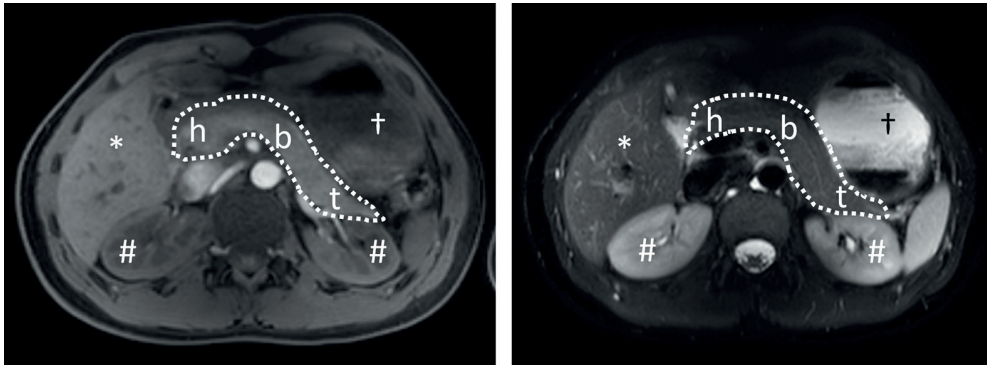


Figure 1.1. Magnetic resonance images (**left:** T1-weighted, **right:** T2-weighted) of the pancreas of a healthy volunteer. The pancreas (dotted lines), including the pancreatic head (h), body (b) and tail (t), the liver (*), the kidneys (#), and the stomach (†) are accentuated.

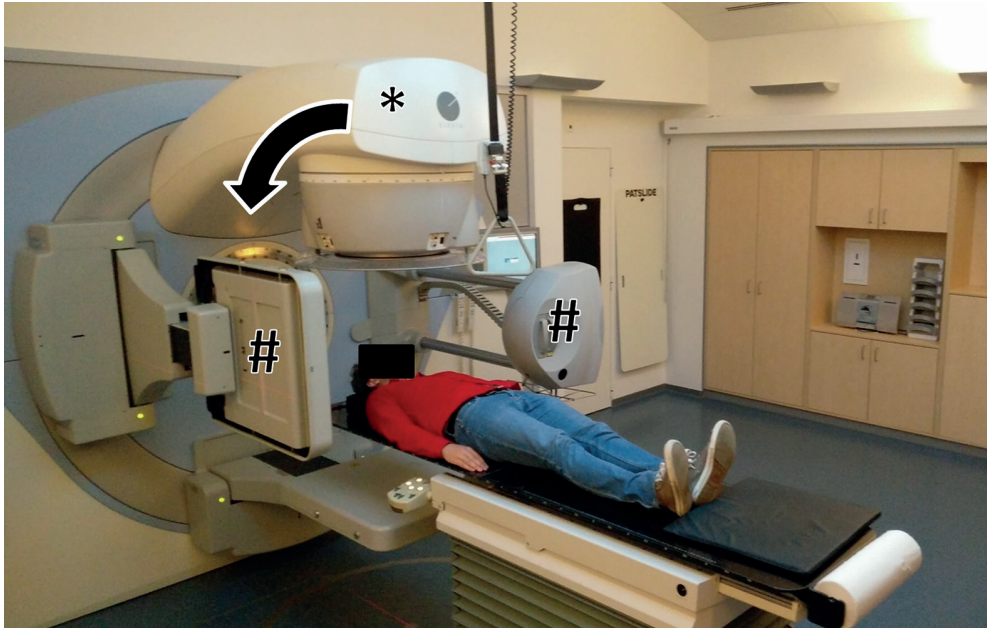


Figure1.2. A photo of a linac. The patient is located at the center of the beam. The linear accelerator gantry (*) from which the radiation beam emerges is able to rotate around the patient (arrow) together with the cone beam computed tomography device (#).

Treatment planning

In order to maximize the dose to the tumor while sparing OARs, it is necessary to image these structures. Therefore, before the start of therapy, computed tomography (CT) images are obtained (Figs. 1.3 and 1.4 a), called a planning CT. A radiation oncologist then delineates the visible tumor and lymph nodes with potential tumor cells on all the image slices of the planning CT. This delineated volume is known as gross tumor volume (GTV; Figs. 1.3 a and c). As microscopic extensions of the tumor may not be visible on the CT images [22], the GTV is expanded by a margin, typically of several millimeters, to generate a clinical treatment volume (CTV). The CTV includes these suspected microscopic extensions. As irradiation often lasts several minutes per session, patients are mostly irradiated during free breathing. To include the CTV during the full respiratory cycle, an additional margin is added to the CTV to generate the internal target volume (ITV). A CTV-ITV margin that incorporates all motion for every patient will overestimate the motion for most individual patients as respiratory motion varies between subjects [23]. Therefore, patient specific margins can be determined instead, using inhalation and exhalation breath-hold CT [24],

4DCT [25] or MRI images [26, 27]. Finally, to deal with all other uncertainties, such as patient setup errors and errors in the original GTV delineation, an additional margin (typically 10–15 mm) is added to the ITV to generate a planning target volume (PTV) [28]. In radiotherapy planning the aim is to develop a plan that irradiates the PTV at a high dose, while minimizing the dose to the OARs. For this reason, the OARs are also delineated on the planning CT.

After generating the PTV and delineating the OARs, a treatment plan is designed. In such a treatment plan, beam settings are optimized such that the predefined required dose to the PTV (tumor) is reached while OARs have an as low dose as possible (Figs. 1.3 b and d). In modern radiotherapy, the angle under which the patient is irradiated can be varied by rotating the beam around the patient. Furthermore, the shape and intensity of the beam can be modulated during treatment. By doing so, very steep dose gradients between the PTV and OARs can be achieved. This is done either using several predefined angles and beam shapes (intensity modulated radiotherapy; IMRT) [29] or a continuously changing beam shape that rotates around the patient (volumetric modulated arc therapy; VMAT) [30].

Treatment

Typically, radiotherapy of pancreatic cancer patients is fractionated over multiple (15–28) daily sessions [31]. During each session, the target volume (i.e. PTV) should be aligned with the radiation beam of the linac to ensure that the planned treatment is delivered accurately. Most linacs contain an integrated cone beam CT (CBCT) device (Fig. 1.2) which allows making 3D volumetric CT-like images of the patient's anatomy while on the treatment couch (Fig. 1.4 b) [32]. Such a CBCT allows determining the shift of the patient's anatomy as presented on the treatment couch compared to the anatomy on the planning CT. The patient's treatment couch can then be moved according to the found shift between both scans to ensure alignment of the PTV with the treatment plan.

Determining the shift between the CT and CBCT is done with respect to some anatomical landmarks visible both on the planning CT and CBCT. As the CBCT has poor soft tissue (e.g., pancreas) contrast, but good contrast of bones, often the bony anatomy was used as landmarks. However, the position of the PTV with respect to the bony anatomy can change between the planning CT and daily CBCT as a result of e.g., bowel filling [33]. Therefore, nowadays small (sub-mm diameter) golden fiducial markers that are visible on CT and CBCT (Fig. 1.4 a-b) are placed endoscopically with thin needles (inner diameter of approximately 0.4 mm) inside the tumor in some clinics

[34-36]. These markers are then used to match the planning CT and CBCT images and ensure alignment of the PTV [33, 37-39].

Finally, respiratory motion during radiotherapy effectively results in blurring of the dose [40]. If ignored, this motion can lead to under treatment of the tumor and additional radiation to OARs. Classically, the respiratory motion is included in the PTV by introducing an ITV, as mentioned above. Including the uncertainties due to respiratory motion results in larger PTVs and hence may cause higher dose to OARs. Therefore, there are several alternative approaches to address the respiratory motion. For example, patients can be treated in mid-ventilation [41] or during breath-holding [42, 43]. Also, gating or tracking techniques may be used [44-46]. These techniques can potentially eliminate the use of an ITV and hence decrease the size of the PTV, reducing toxicity.

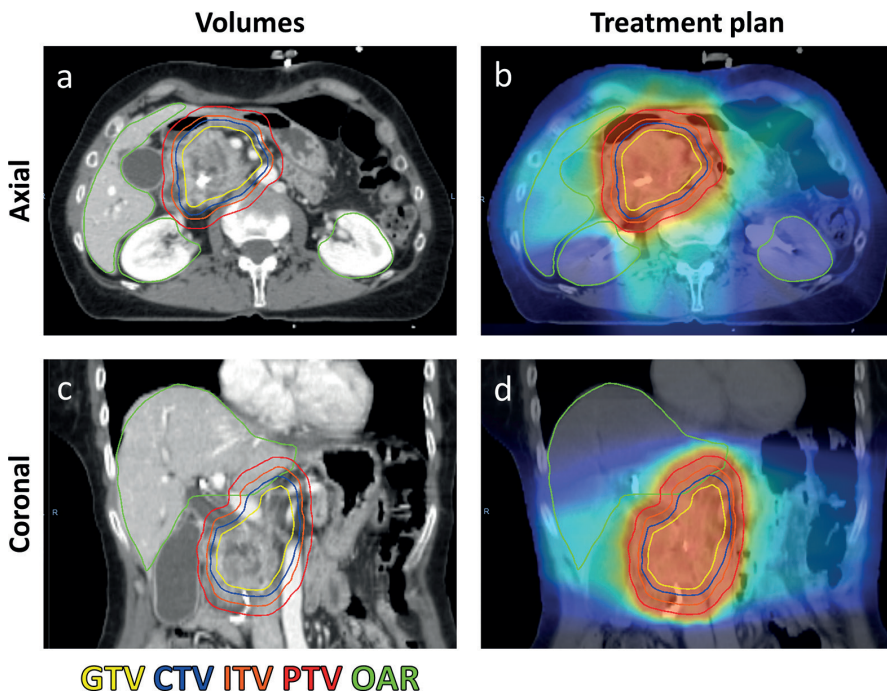


Figure 1.3. Delineated GTV (yellow), CTV (blue), ITV (orange) and PTV (red) volumes, as well as some OARs (liver + kidneys; green) projected on a planning CT scan in axial (a) and coronal (c) view. Treatment plans based on these delineated volumes are presented too (b, d).

There is further room for improvement in radiotherapy of pancreatic cancer patients. Currently, there is a large variation between observers in GTV delineation (i.e. a poor interobserver agreement) [47-49]. As the entire treatment is based on the delineated GTV, any error in this delineation impacts the entire treatment. Therefore, GTV delineations should be as accurate as possible. These delineations are currently often based on CT images, which have poor soft tissue contrast and hence poor pancreatic tissue and tumorous tissue contrast (Fig. 1.4 a-b). Potentially scans with better soft tissue contrast may improve the accuracy of the GTV delineation. Furthermore, some patients have pancreatic cancers that are unresponsive to radiotherapy. For these patients, radiotherapy is undesirable as the irradiation will only damage healthy tissue. As discussed further on, both issues can potentially be addressed by MRI.

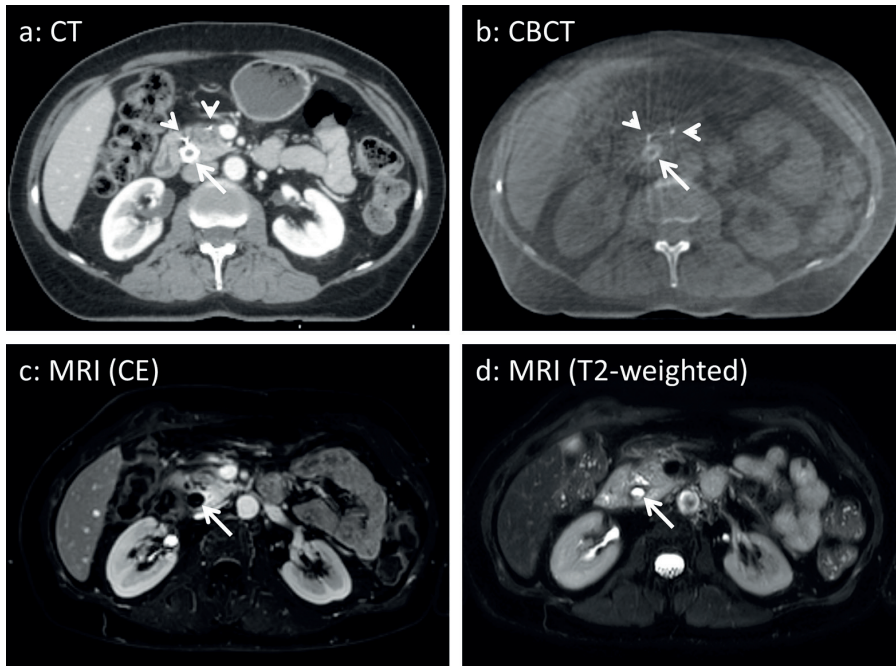


Figure 1.4. CT (a), CBCT (b) and MRI (c-d) images of a pancreatic cancer patient. The biliary stent (arrow) and fiducial markers in the pancreatic tumor (arrowhead, not visible in MRI images) are accentuated.

Magnetic resonance imaging

MRI is a medical imaging technique that allows for imaging patient's anatomy. MRI manipulates proton spins of specific atoms (typically hydrogen) using strong magnetic fields (typically 1.5 T to 3 T) and a sequence of radiofrequency pulses and magnetic field gradients, called pulse sequence, to generate images [50]. By varying settings of the pulse sequence, different types of contrasts can be achieved, such as T1- or T2-weighted contrasts (e.g., Fig. 1.5 a-b). Furthermore, MRI can be used to quantify specific tissue properties, such as the T1-values [51], T2-values [52, 53], diffusivity [54] and the amount of perfusion [55, 56] of tissue. In such a case, maps are created of these specific tissue properties (e.g., Fig. 1.5 d-f).

In this thesis we focus on the following sequences: T1-weighted, contrast enhanced (CE) T1-weighted, T2-weighted, and DWI images (Fig. 1.5). T1-weighted images of the pancreas can be obtained in a single breath-hold at high resolution [57]. However, most tumors, are poorly visible on plain T1-weighted images and therefore the acquisition is repeated after an intravenous gadolinium contrast injection to generate CE images. As pancreatic tumors are often poorly perfused, the contrast medium will initially mainly enhance the signal from healthy tissue, resulting in hypointense (i.e. darker) tumor tissue compared to surrounding more enhanced healthy tissue in such CE images (Fig. 1.5 c). T2-weighted images often have high in-plane resolutions, but poor through-plane resolution [57]. Furthermore, T2-weighted images are often acquired during several breath-holds, which can result in a mismatch of anatomy between neighboring slices. T2-weighted images have good soft tissue contrast for OARs. Moreover, T2-weighted images offer complementary information to T1-weighted images which can help in differentiating between tumor, necrosis and inflammation.

Diffusion-weighted imaging

In DWI, gradients in the magnetic field (called diffusion gradients) are applied before signal read-out to sensitize the MRI signal to the diffusion of water molecules. Diffusion is the intermingling of molecules as a result of random motion of the molecules due to their kinetic energy. DWI uses the diffusion of hydrogen molecules to generate contrast in images. In DWI, images with different diffusion weightings are acquired, which is achieved by varying the diffusion gradients' strength. A 4D data set is then created, in which the signal is described as function of diffusion weighting for the three spatial dimensions. For each voxel in the spatial dimensions, a diffusion model can be fitted to the signal as function of diffusion weighting. In the classical diffusion model, the signal attenuation as a function of diffusion weighting is modeled

monoexponentially. The model parameter apparent diffusion coefficient (ADC) is then found by fitting this monoexponential DWI model. When such a monoexponential fit is made for each voxel, an ADC-map can be generated (Fig. 1.5 d). ADC represents the diffusivity of water in tissue. The ADC depends on the microscopic structure of tissue and often gives good contrast between pancreatic tumors and healthy tissue [58-61].

However, many competing models describe DWI data. The most well-known competing model is the intravoxel incoherent motion (IVIM) model (Fig. 1.5 e-f). In the IVIM model, the sensitivity of DWI-signal to perfusion effects is modeled as well. Perfusion is the flow of blood through tissue. In this model, the signal from blood in capillaries dephases more rapidly as a function of diffusion weighting than signal from tissue. To model both components, the IVIM model assumes DWI data decays bi-exponentially as a function of diffusion weighting. Since the introduction of the IVIM model, the non-monoexponential behavior of DWI data in the pancreas was confirmed in multiple studies [60-65] and related to perfusion [66, 67]. The added perfusion parameters of the IVIM model have shown additional value for lesion characterization in the pancreas [60-65] and enabled treatment response monitoring in various other organs [68, 69].

DWI is often limited to poor resolutions and acquisition times of several minutes. Therefore, DWI is either imaged during free breathing [70, 71], in multiple breath-holds [62] or during respiratory gating/triggering [72].

MRI and radiotherapy

In this thesis two potential applications of MRI for radiotherapy are discussed: MRI during treatment planning and MRI for monitoring treatment response. For several organs, it was shown that when MRI, which has good soft tissue contrast (Fig. 1.5 c-d) [57, 73-75], is added during treatment planning, delineations improved [76-79]. Therefore, MRI may also improve tumor delineation for pancreatic cancer patients. Andreychenko et al. [80] suggested that the most useful MRI images for target volume delineation of pancreatic cancer patients would be T1-weighted, contrast enhanced (CE), T2-weighted, and DWI images (Fig. 1.5 a-d). Whether GTV delineation is improved with MRI has not been investigated for pancreatic cancer patients.

Furthermore, some tumors are insensitive to irradiation, so-called radioresistant tumors. If the lack of treatment response could be detected at an early stage of treatment (known as treatment response monitoring), or, ultimately, if tumor sensitivity to irradiation can be determined before treatment, unnecessary irradiation can be prevented. There are several promising MRI pulse sequences that may be able

to monitor the response of the tumor to treatment or determine the tumor sensitivity to irradiation before treatment. Part of this thesis focusses on one of these pulse sequences, namely DWI and in particular the IVIM model for DWI. It is known that low diffusivity in the tumor before treatment is related to poor treatment outcome in both chemotherapy [81] and radiochemotherapy [82] of the pancreas. It was suggested that low perfusion before treatment of several therapies may also relate to limited treatment response [83]. Therefore, IVIM modeling of DWI may prove a useful tool for treatment response monitoring and treatment outcome prediction in the pancreas.

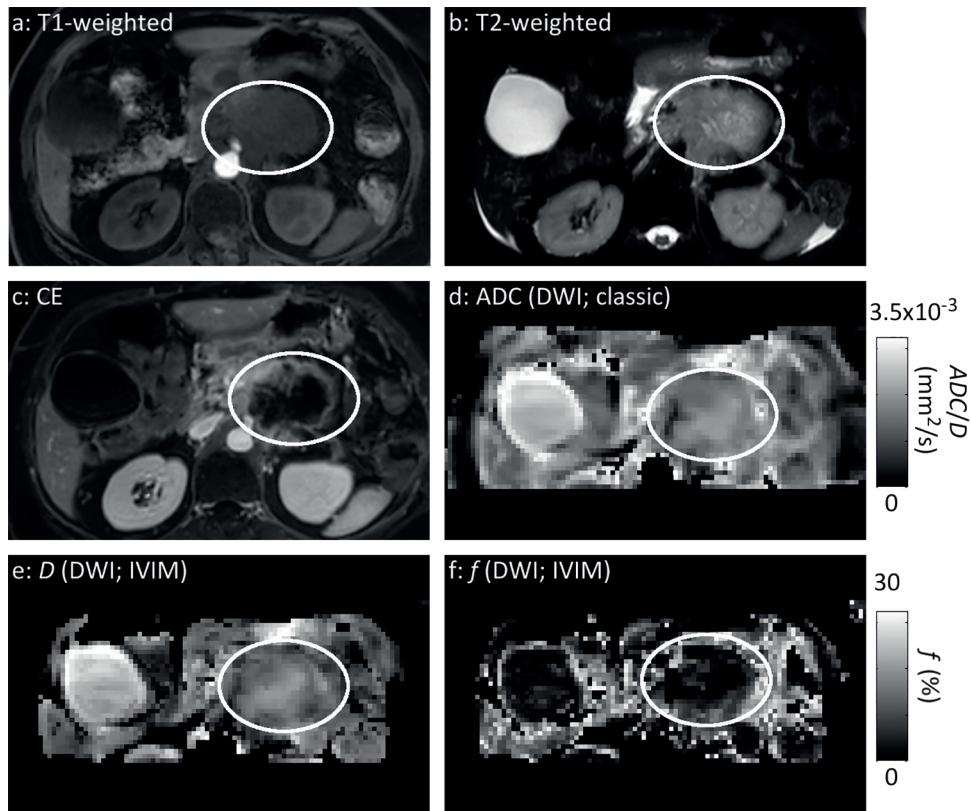


Figure 1.5. Six different MRI images of the same patient. Images include standard non-CE (a, b), CE T1-weighted (c) and DWI (d-f) images. Diffusion maps obtained from different fit models (d, e) and a perfusion fraction map (f) are displayed. The white ellipse surrounds the tumor.

Outline of this thesis

In this thesis, MRI techniques are optimized for the purpose of radiotherapy of pancreatic cancer patients, and the added value of MRI is assessed.

First, in **chapter 2** the interobserver variation of tumor delineation on CT images is quantified to assess the current clinical practice as a baseline. Then, in **chapter 3**, the value of offering MRI images alongside the CT images for delineating pancreatic cancers is quantified.

When MRI is offered alongside CT, it is desirable that the images are matched, which is best done using intratumoral fiducial markers [33, 37]. Therefore, in **chapter 4** the visibility and artifacts caused by fiducial markers on MRI images are quantified using a pulse sequence-independent approach. As 70% of the pancreatic cancer patients receive biliary stents, which are often located close to the tumor, the same approach is used to quantify artifacts caused by biliary stents in **chapter 5**.

During radiotherapy, tumors are treated with high precision and hence 3D high-resolution imaging is desired. A disadvantage of common T2-weighted pulse sequences is its thick slices. Therefore, in **chapter 6**, an alternative T2-weighted-like pulse sequence that allows for high 3D resolution is optimized for imaging of the pancreas.

In **chapter 7** the acquisition of DWI for IVIM model fitting in the pancreas and liver is optimized. In **chapter 8** the fit algorithm for the IVIM model is optimized, and the IVIM model is compared to other DWI data models considering pancreatic tumor contrast, treatment monitoring, and treatment outcome prediction.

Finally, in **chapter 9** the findings of this thesis are set in clinical perspective and the future of MRI for pancreatic radiotherapy is discussed.

References

1. Bilimoria KY, Bentrem DJ, Ko CY et al. Validation of the 6th edition AJCC pancreatic cancer staging system: Report from the National Cancer Database. *Cancer* 2007; 110(4):738–744.
2. Sharma C, Eltawil KM, Renfrew PD et al. Advances in diagnosis, treatment and palliation of pancreatic carcinoma: 1990–2010. *World J. Gastroenterol.* 2011; 17(7):867–897.
3. Ferlay J, Soerjomataram I, Dikshit R et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 2015; 136(5):E359–E386.
4. Nitecki SS, Sarr MG, Colby T V, van Heerden JA. Long-term survival after resection for ductal adenocarcinoma of the pancreas. Is it really improving? *Ann. Surg.* 1995; 221(1):59–66.
5. Ferlay JPCBF. More deaths from pancreatic cancer than breast cancer in the EU by 2017. *Acta Oncol* 2016; 23(9–10):1158–1160.
6. Whipple AO, Parsons WB, Mullins CR. Treatment of carcinoma of the ampulla of Vater. *Ann. Surg.* 1935; 102(4):763–779.
7. Katz MHG, Hwang R, Fleming JB, Evans DB. Tumor-Node-Metastasis Staging of Pancreatic Adenocarcinoma. *CA. Cancer J. Clin.* 2008; 58(2):111–125.
8. Wong JC, Raman S. Surgical resectability of pancreatic adenocarcinoma: CTA. *Abdom. Imaging* 2010; 35(4):471–480.
9. Ryan DP, Hong TS, Bardeesy N. Pancreatic Adenocarcinoma. *N. Engl. J. Med.* 2015; 371(11):1039–1049.
10. Konstantinidis IT, Warshaw AL, Allen JN et al. Pancreatic ductal adenocarcinoma: is there a survival difference for R1 resections versus locally advanced unresectable tumors? What is a “true” R0 resection? *Ann. Surg.* 2013; 257(4):731–736.
11. Sohal DPS, Mangu PB, Khorana AA et al. Metastatic Pancreatic Cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J. Clin. Oncol.* 2016; 34(23):2784–2796.
12. Seufferlein T, Bachet JB, Van Cutsem E, Rougier P. Pancreatic adenocarcinoma: ESMO-ESDO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 2012; 23(SUPPL. 7):vi33–40.
13. Klinkenbijn JH, Jeekel J, Sahmoud T et al. Adjuvant radiotherapy and 5-fluorouracil after curative resection of cancer of the pancreas and periampullary region: phase III trial of the EORTC gastrointestinal tract cancer cooperative group. *Ann. Surg.* 1999; 230(6):776–784.
14. Neoptolemos JP, Stocken DD, Bassi C et al. Adjuvant chemotherapy with fluorouracil plus folinic acid vs gemcitabine following pancreatic cancer resection: a randomized controlled trial. *JAMA* 2010; 304(10):1073–1081.
15. Neoptolemos JP, Stocken DD, Friess H et al. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. *N. Engl. J. Med.* 2004; 350(12):1200–1210.
16. Oettle H, Post S, Neuhaus P et al. Adjuvant chemotherapy with gemcitabine vs observation in patients undergoing curative-intent resection of pancreatic cancer: a randomized controlled trial. *JAMA* 2007; 297(3):267–277.
17. Takahashi H, Ohigashi H, Gotoh K et al. Preoperative gemcitabine-based chemoradiation therapy for resectable and borderline resectable pancreatic cancer. *Ann. Surg.* 2013; 258(6):1040–1050.
18. Kim EJ, Ben-Josef E, Herman JM et al. A multi-institutional phase 2 study of neoadjuvant gemcitabine and oxaliplatin with radiation therapy in patients with pancreatic cancer. *Cancer* 2013; 119(15):2692–2700.
19. Greer SE, Pipas JM, Sutton JE et al. Effect of Neoadjuvant Therapy on Local Recurrence after Resection of Pancreatic Adenocarcinoma. *J. Am. Coll. Surg.* 2008; 206(3):451–457.
20. Versteijne E, van Eijck CHJ, Punt CJA et al. Preoperative radiochemotherapy versus immediate surgery for resectable and borderline resectable pancreatic cancer (PREOPANC trial): study protocol for a multicentre randomized controlled trial. *Trials* 2016; 17(1):127.

21. Evans DB, Varadhachary GR, Crane CH et al. Preoperative gemcitabine-based chemoradiation for patients with resectable adenocarcinoma of the pancreatic head. *J. Clin. Oncol.* 2008; 26(21):3496–3502.
22. Arvold ND, Niemierko A, Mamon HJ et al. Pancreatic cancer tumor size on CT scan versus pathologic specimen: Implications for radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.* 2011; 80(5):1383–1390.
23. Feng M, Balter JM, Normolle D et al. Characterization of Pancreatic Tumor Motion Using Cine MRI: Surrogates for Tumor Position Should Be Used With Caution. *Int. J. Radiat. Oncol. Biol. Phys.* 2009; 74(3):884–891.
24. Yang W, Fraass BA, Reznik R et al. Adequacy of inhale/exhale breathhold CT based ITV margins and image-guided registration for free-breathing pancreas and liver SBRT. *Radiat. Oncol.* 2014; 9:11.
25. Hallman JL, Mori S, Sharp GC et al. A four-dimensional computed tomography analysis of multiorgan abdominal motion. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 83(1):435–441.
26. Stemkens B, Tijssen RHN, De Senneville BD et al. Optimizing 4-dimensional magnetic resonance imaging data sampling for respiratory motion analysis of pancreatic tumors. *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 91(3):571–578.
27. Tryggestad E, Flammang A, Han-Oh S et al. Respiration-based sorting of dynamic MRI to derive representative 4D-MRI for radiotherapy planning. *Med. Phys.* 2013; 40(5):51909.
28. Stroom JC, Heijmen BJM. Geometrical uncertainties, radiotherapy planning margins, and the ICRU-62 report. *Radiother. Oncol.* 2002; 64(1):75–83.
29. Intensity Modulated Radiation Therapy Collaborative Working Group –. Intensity-modulated radiotherapy: current status and issues of interest. *Int. J. Radiat. Oncol. Biol. Phys.* 2001; 51(4):880–914.
30. Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med. Phys.* 2008; 35(1):310–317.
31. Wang F, Kumar P. The role of radiotherapy in management of pancreatic cancer. *J. Gastrointest. Oncol.* 2011; 2(3):157–167.
32. Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2002; 53(5):1337–1349.
33. van der Horst A, Wognum S, Dávila Fajardo R et al. Interfractional position variation of pancreatic tumors quantified using intratumoral fiducial markers and daily cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 87(1):202–208.
34. Dávila Fajardo R, Lekkerkerker SJ, van der Horst A et al. EUS-guided fiducial markers placement with a 22-gauge needle for image-guided radiation therapy in pancreatic cancer. *Gastrointest. Endosc.* 2014; 79(5):851–855.
35. Varadarajulu S, Bang JY, Hebert-Magee S. Assessment of the technical performance of the flexible 19-gauge EUS-FNA needle. *Gastrointest. Endosc.* 2012; 76(2):336–343.
36. Dehnad H, Nederveen AJ, van der Heide UA et al. Clinical feasibility study for the use of implanted gold seeds in the prostate as reliable positioning markers during megavoltage irradiation. *Radiother. Oncol.* 2003; 67(3):295–302.
37. van der Horst A, Lens E, Wognum S et al. Limited role for biliary stent as surrogate fiducial marker in pancreatic cancer: Stent and intratumoral fiducials compared. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):641–648.
38. Nederveen A, Lagendijk J, Hofman P. Detection of fiducial gold markers for automatic on-line megavoltage position verification using a marker extraction kernel (MEK). *Int. J. Radiat. Oncol. Biol. Phys.* 2000; 47(5):1435–1442.
39. Nederveen AJ, Dehnad H, van der Heide UA et al. Comparison of megavoltage position verification for prostate irradiation based on bony anatomy and implanted fiducials. *Radiother. Oncol.* 2003; 68(1):81–88.

40. Gierga DP, Chen GTY, Kung JH et al. Quantification of respiration-induced abdominal tumor motion and its impact on IMRT dose distributions. *Int. J. Radiat. Oncol. Biol. Phys.* 2004; 58(5):1584–1595.
41. Lens E, van der Horst A, Versteijne E et al. Dosimetric advantages of midventilation compared with internal target volume for radiation therapy of pancreatic cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 92(3):675–682.
42. Lens E, Gurney-Champion OJ, Tekelenburg DR et al. Abdominal organ motion during inhalation and exhalation breath-holds: pancreatic motion at different lung volumes compared. *Radiother. Oncol.* 2016. In press
43. Lens E, Gurney-Champion O, van der Horst A et al. MO-FG-BRA-09: Towards an optimal breath-holding procedure for radiotherapy: Differences in organ motion during inhalation and exhalation breath-holds. *Med. Phys.* 2016; 43(6):3711.
44. Shirato H, Shimizu S, Kitamura K, Onimaru R. Organ motion in image-guided radiotherapy: Lessons from real-time tumor-tracking radiotherapy. *Int. J. Clin. Oncol.* 2007; 12(1):8–16.
45. Verellen D, Depuydt T, Gevaert T et al. Gating et tracking 4D dans les tumeurs thoraciques. *Cancer/ Radiothérapie* 2010; 14(6–7):446–454.
46. Huguet F, Yorke ED, Davidson M et al. Modeling Pancreatic Tumor Motion Using 4-Dimensional Computed Tomography and Surrogate Markers. *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 91(3):579–587.
47. Caravatta L, Macchia G, Mattiucci GC et al. Inter-observer variability of clinical target volume delineation in radiotherapy treatment of pancreatic cancer: A multi-institutional contouring experience. *Radiat. Oncol.* 2014; 9(1):198.
48. Yamazaki H, Nishiyama K, Tanaka E et al. Dummy run for a phase II multi-institute trial of chemoradiotherapy for unresectable pancreatic cancer: Inter-observer variance in contour delineation. *Anticancer Res.* 2007; 27(4 C):2965–2971.
49. Fokas E, Clifford C, Spezi E et al. Comparison of investigator-delineated gross tumor volumes and quality assurance in pancreatic cancer: Analysis of the pretrial benchmark case for the SCALOP trial. *Radiother. Oncol.* 2015; 117(3):432–437.
50. Lauterbur PC. Image formation by induced local interactions. Examples employing nuclear magnetic resonance. *Nat. (London, United Kingdom)* 1973; 242(5394):190–191.
51. Look DC, Locker DR. Time saving in measurement of NMR and EPR relaxation times. *Rev. Sci. Instrum.* 1970; 41(2):250–251.
52. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* 1954; 94(3):630–638.
53. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* 1958; 29(8):688–691.
54. Le Bihan D, Breton E, Lallemand D et al. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 1986; 161(2):401–407.
55. Petersen ET, Zimine I, Ho Y-CL, Golay X. Non-invasive measurement of perfusion: a critical review of arterial spin labelling techniques. *Br. J. Radiol.* 2006; 79(944):688–701.
56. Le Bihan D, Breton E, Lallemand D et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 1988; 168(2):497–505.
57. Sandrasegaran K, Lin C, Akisik FM, Tann M. State-of-the-art pancreatic MRI. *Am. J. Roentgenol.* 2010; 195(1):42–53.
58. Matsuki M, Inada Y, Nakai G et al. Diffusion-weighted MR imaging of pancreatic carcinoma. *Abdom. Imaging* 2007; 32(4):481–483.
59. Barral M, Taouli B, Guio B et al. Diffusion-weighted MR imaging of the pancreas: Current status and recommendations. *Radiology* 2015; 274(1):45–63.
60. Lemke A, Laun FB, Klauf M et al. Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: Comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. *Invest. Radiol.* 2009; 44(12):769–775.

61. Re TJ, Lemke A, Klaub M et al. Enhancing pancreatic adenocarcinoma delineation in diffusion derived intravoxel incoherent motion f-maps through automatic vessel and duct segmentation. *Magn. Reson. Med.* 2011; 66(5):1327–1332.
62. Klaub M, Lemke A, Grünberg K et al. Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. *Invest. Radiol.* 2011; 46(1):57–63.
63. Kang KM, Lee JM, Yoon JH et al. Intravoxel incoherent motion diffusion-weighted MR imaging for characterization of focal pancreatic lesions. *Radiology* 2014; 270(2):444–453.
64. Hwang EJ, Lee JM, Yoon JH et al. Intravoxel incoherent motion diffusion-weighted imaging of pancreatic neuroendocrine tumors: prediction of the histologic grade using pure diffusion coefficient and tumor size. *Invest. Radiol.* 2014; 49(6):396–402.
65. Klaub M, Mayer P, Bergmann F et al. Correlation of histological vessel characteristics and diffusion-weighted imaging intravoxel incoherent motion derived parameters in pancreatic ductal adenocarcinomas and pancreatic neuroendocrine tumors. *Invest. Radiol.* 2015; 50(11):792–797.
66. Heusch P, Wittsack HJ, Heusner T et al. Correlation of biexponential diffusion parameters with arterial spin-labeling perfusion MRI: Results in transplanted kidneys. *Invest. Radiol.* 2013; 48(3):140–144.
67. Lemke A, Laun FB, Simon D et al. An in vivo verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. *Magn. Reson. Med.* 2010; 64(6):1580–1585.
68. Gaeta M, Benedetto C, Minutoli F et al. Use of diffusion-weighted, intravoxel incoherent motion, and dynamic contrast-enhanced MR imaging in the assessment of response to radiotherapy of lytic bone metastases from breast cancer. *Acad. Radiol.* 2014; 21(10):1286–1293.
69. Koh DM. Science to practice: Can intravoxel incoherent motion diffusion-weighted MR imaging be used to assess tumor response to antivascular drugs? *Radiology* 2014; 272(2):307–308.
70. Leporq B, Saint-Jalmes H, Rabrait C et al. Optimization of intra-voxel incoherent motion imaging at 3.0 Tesla for fast liver examination. *J. Magn. Reson. Imaging* 2014; 0:1–9.
71. Schoennagel BP, Habermann CR, Roesch M et al. Diffusion-weighted imaging of the healthy pancreas: apparent diffusion coefficient values of the normal head, body, and tail calculated from different sets of b-values. *J. Magn. Reson. Imaging* 2011; 34(4):861–865.
72. Ma C, Liu L, Li YJ et al. Intravoxel incoherent motion MRI of the healthy pancreas: Monoexponential and biexponential apparent diffusion parameters of the normal head, body and tail. *J. Magn. Reson. Imaging* 2015; 41(5):1236–1241.
73. Schima W. MRI of the pancreas: tumours and tumour-simulating processes. *Cancer Imaging* 2006; 6:199–203.
74. Erturk SM, Alberich-Bayarri A, Herrmann KA et al. Use of 3.0-T MR Imaging for Evaluation of the Abdomen. *Radiographics* 2009; 29(6):1547–1563.
75. Park HS, Lee JM, Choi HK et al. Preoperative evaluation of pancreatic cancer: Comparison of gadolinium-enhanced dynamic MRI with MR cholangiopancreatography versus MDCT. *J. Magn. Reson. Imaging* 2009; 30(3):586–595.
76. Rasch C, Keus R, Pameijer FA et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 1997; 39(4):841–848.
77. Villeirs GM, van Vaerenbergh K, Vakaet L et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther. Onkol.* 2005; 181(7):424–430.
78. Rasch C, Barillot I, Remeijer P et al. Definition of the prostate in CT and MRI: A multi-observer study. *Int. J. Radiat. Oncol. Biol. Phys.* 1999; 43(1):57–66.
79. den Hartogh MD, Philippens MEP, van Dam IE et al. MRI and CT imaging for preoperative target volume delineation in breast-conserving therapy. *Radiat. Oncol.* 2014; 9(1):63.
80. Andreychenko A, Heerkens H, Meijer G et al. SU-E-J-07: A Functional MR Protocol for the Pancreatic Tumor Delineation. *Med. Phys.* 2014; 41(6):155–156.

81. Niwa T, Ueno M, Ohkawa S et al. Advanced pancreatic cancer: the use of the apparent diffusion coefficient to predict response to chemotherapy. *Br. J. Radiol.* 2009; 82(973):28–34.
82. Cuneo KC, Chenevert TL, Ben-Josef E et al. A pilot study of diffusion-weighted MRI in patients undergoing neoadjuvant chemoradiation for pancreatic cancer. *Transl. Oncol.* 2014; 7(5):644–649.
83. Moulder JE, Rockwell S. Tumor hypoxia: its impact on cancer therapy. *Cancer Metastasis Rev.* 1987; 5(4):313–341.



CHAPTER 2

Delineation on CT

Considerable interobserver variation in delineation of pancreatic cancer on 3DCT and 4DCT: a multi-institutional study

Eva Versteijne*
Oliver J Gurney-Champion*
Astrid van der Horst
Eelco Lens
M Willemijn Kolff
Jeroen Buijsen
Gati Ebrahimi
Karen J Neelis
Coen RN Rasch
Jaap Stoker
Marcel van Herk
Arjan Bel
Geertjan van Tienhoven

*both authors contributed equally to this work

submitted

Abstract

Background: The delineation of pancreatic tumors on CT is challenging. In this study, we quantified the interobserver variation for pancreatic tumor delineation on 3DCT as well as on 4DCT.

Methods: Eight observers (radiation oncologists) from six institutions delineated pancreatic tumors of four patients with (borderline) resectable pancreatic cancer. The study consisted of two stages. In the 3DCT-stage, the gross tumor volume (GTV) was delineated on a contrast-enhanced scan. In the 4DCT-stage, the internal GTV (iGTV) was delineated, accounting for the respiratory motion. We calculated the volumes of the (i)GTV, the overlap of the delineated volumes (expressed as generalized conformity index: CI_{gen}), the local observer variation (local standard deviation: SD) and the overall observer variation (overall SD). We compared these results between GTVs and iGTVs. Additionally, observers filled out a questionnaire concerning the difficulty of the delineation and their experience in delineating pancreatic tumors.

Results: The ratios of the largest to the smallest delineated GTV and iGTV within the same patient were 6.8 and 16.5, respectively. As the iGTV incorporates the GTV during all respiratory phases, the mean volumes of the iGTV (40.07 cm³) were larger than those of the GTV (29.91 cm³). For all patients, CI_{gen} was larger for the iGTV than for the GTV. The mean overall observer variation (root-mean-square of all local SD s over four patients) was 0.63 cm and 0.80 cm for the GTV and iGTV, respectively. The largest local observer variations were seen close to biliary stents and suspicious pathological enlarged lymph nodes, as some observers included them and some did not. This variation was more pronounced for the iGTV than for the GTV.

The observers rated the 3DCT-stage and 4DCT-stage equally difficult and treated on average three to four pancreatic cancer patients per year.

Conclusions: A considerable interobserver variation in delineation of pancreatic tumors was observed. This variation was larger for 4D than for 3D delineation. The largest local observer variation was found around biliary stents and suspicious pathological enlarged lymph nodes.

Background

The aim of radiotherapy is delivering a high radiation dose to the tumor while minimizing the dose to organs at risk (OARs). For pancreatic tumors, this is challenging due to day-to-day position variation, respiratory motion, and uncertainties in delineation of the tumor [14].

Radiation oncologist delineate the gross tumor volume (GTV) on a three-dimensional CT (3DCT). The GTV is expanded with a margin to account for microscopic extensions, resulting in the clinical target volume (CTV). For the remaining uncertainties, such as internal and set-up uncertainties, an additional margin is added to form the planning target volume (PTV). Nowadays, a four-dimensional CT (4DCT) scan is increasingly used to account for tumor motion during respiration [1, 2], for example combined with the internal target volume (ITV) [5] or mid-ventilation approach [6]. For pancreatic cancer patients treated at our department, we combine 4DCT with a modified ITV approach. In this approach, the radiation oncologist delineates the GTV on the average scan of the 4DCT and expands that on all respiratory phases of the 4DCT to generate an internal GTV (iGTV). A 5 mm margin is then added to define the internal CTV (iCTV). An additional PTV margin is added, to account for remaining set-up uncertainties. This PTV margin can be smaller compared with 3DCT delineation since respiratory motion uncertainty is accounted for in this 4D approach. In both the 3DCT and 4DCT approaches it is important that the appropriate margin size is used as too small a margin leads to under-treatment of the target volume whereas too large a margin leads to unnecessarily high doses to the OARs. The PTV margins currently used to account for the delineation uncertainties in pancreatic cancer are largely based on estimates of these uncertainties. To investigate whether these estimates are correct we performed a delineation study.

Previous delineation studies quantified the interobserver delineation uncertainties for several tumor sites [7-11]. These studies resulted in standardized delineation protocols for those organs. For pancreatic cancer, such a protocol is only available in the postoperative setting [12]. Only three multi-institutional studies on the delineation of pancreatic tumors are available [13-15]. These studies show large interobserver variation in GTV delineation, with ratios of the largest to the smallest GTV volume of 6.8 [13], 9 [14] and 3 [15]. Two of these studies [14, 15] were quality control studies of a clinical trial. Those studies only used 3DCT and included 1–2 patients with locally advanced pancreatic cancer. The third [13], was a delineation study which included two patients and only investigated the interobserver variation using 3DCT. All of these studies [13-15] reported limited quantitative information (i.e. standard deviation, SD and generalized conformity index, CI_{gen}).

The aim of this study was to quantify the interobserver variation for GTV (using 3DCT) and iGTV (using 4DCT) delineations. The study included four patients with (borderline) resectable pancreatic cancer, and eight radiation oncologists from six institutions.

Methods

Radiation oncologists (observers) from all nine institutions participating in the PREOPANC trial were asked to participate in this delineation study. Eight observers from six institutions actually participated.

Patients' characteristics

The data of four patients with histologically proven (borderline) resectable pancreatic tumors were used and anonymized. All patients gave written informed consent for both the PREOPANC trial (EudraCT number 2012-003181-40) and MIPA study (NCT01989000) and were the first four patients that randomized for preoperative radiochemotherapy at the Academic Medical Center (AMC) within the PREOPANC trial [16]. Both studies were approved by the local medical ethics committee (PREOPANC: Erasmus Medical Center, Rotterdam; MIPA: AMC, Amsterdam) [16]. Preoperative radiochemotherapy consisted of 15 fractions of 2.4 Gy combined with gemcitabine 1000 mg/m² once a week for three weeks, preceded and followed by a modified course of gemcitabine 1000 mg/m², once a week for two weeks. Between the three cycles there was one week rest [16].

CT scans

All patients had a contrast-enhanced diagnostic CT scan in the referring hospital, which was considered to be of adequate diagnostic quality by abdominal radiologists from the AMC with extensive experience in pancreatic cancer. The scans included an axial scan in arterial contrast phase (on average 35 seconds after injection, all patients), venous contrast phase (on average 60 seconds after injection, patients 1,2 and 4) and/or a portal contrast phase (on average 240 seconds after injection, patients 1 and 4) with or without reconstructed coronal views. Two experienced radiologists from the AMC reported the studies. The report of patient 2 described two suspicious loco regional lymph nodes; the report of patient 4 described some (cited in the radiology report) enlarged lymph nodes, which were not further characterized.

The planning CT scans were obtained at the radiation oncology department of the AMC with a GE LightSpeed RT 16 scanner (General Electric Company, Waukesha, WI) using a standard acquisition protocol (slice thickness of 2.5 mm). Patients were scanned in treatment position: supine on a flat table top with arms raised above their heads.

First, a 3DCT scan was obtained during free breathing after intravenous Iodine contrast injection. During the same CT session, a few minutes after the 3DCT scan, a 4DCT scan was obtained. The patient's breathing motion was monitored and synchronized to the CT acquisition by the respiratory gating system RPM (Real-Time Position Management, Varian Oncology Systems, Palo Alto, CA). For the 4DCT, images were captured during continuous respiration and divided into ten respiratory bins, resulting in ten image sets of the respiratory cycle. Also, a maximum intensity projection (MIP) and an average intensity projection (Ave-IP) were reconstructed from the ten phase scans. The planning CT scan was obtained during the first modified course of gemcitabine (on average eight days after the first administration of gemcitabine), and on average six weeks (46–62 days, with a mean of 53 days) after the diagnostic CT. The 3DCT and 4DCT scans were registered to each other but not to the diagnostic CT scans.

Fiducial markers and biliary stents

All four patients had intratumoral fiducial markers, which were placed under the guidance of endoscopic ultrasound (EUS), for position verification during radiotherapy [17, 18]. Patients 1, 3 and 4 had a pancreatic head tumor and had received three intratumoral Visicoil fiducial markers (RadioMed, Barlett, TN). For patient 2, two Gold Anchor fiducial markers (Naslund Medical AB, Huddinge, Sweden) and one Visicoil fiducial marker had mistakenly been placed in the pancreas head instead of in the corpus tumor. Also, all patients had biliary drainage: patients 1–3 had fully covered metal biliary stents, patient 4 had external percutaneous biliary drainage. All markers, biliary stents and percutaneous biliary drainage had been placed after the diagnostic CT scans and were thus only visible on the planning CT scan.

Delineation software

The Big Brother software, dedicated to recording delineations as well as observer-computer interactions for radiotherapy delineation studies, was used [8]. Each observer received a USB stick containing all CT scans, the radiology report, the Big

Brother software, and delineation instructions. These instructions were identical to those in the PREOPANC trial protocol [16].

Delineation protocol

The study consisted of a 3DCT-stage and a 4DCT-stage.

In the 3DCT-stage, the observers were asked to delineate the GTV on the 3DCT scan, which was displayed on the main window. The GTV was defined as the macroscopically visible tumor and neighboring suspicious pathological lymph nodes. A separate window was available for viewing the diagnostic CT scans. A margin of 5 mm was automatically applied to create the CTV.

In the 4DCT-stage, the Ave-IP reconstruction was displayed in the main window. The observers were asked to delineate the GTV on the Ave-IP reconstruction and then create an iGTV defined as the volume encompassing the GTV on all ten respiratory phase image sets of the 4DCT. The diagnostic CT scan, 3DCT scan, and remaining 4DCT images including the MIP reconstruction were available in a separate window. As the 3DCT and 4DCT scans from the planning CT were obtained in the same session, the 3DCT and 4DCT scans were linked to the Ave-IP reconstruction displayed in the main window. Furthermore, a copy of the cursor was displayed at the corresponding location in the secondary window when these scans were displayed. Once finished with the iGTV delineations, a margin of 5 mm was automatically applied to create the ICTV. Completed delineations were sent back to the investigators by email.

Questionnaire

Observers were asked to fill out a questionnaire containing eight questions about the delineation process (Additional file 1). These multiple choice questions about the delineation process included answers ranging from very easy to very difficult in five steps. In addition, there were three questions about the experience of the observer in delineating pancreatic tumors as well as the number of pancreatic cancer patients the observers treated yearly within and outside the PREOPANC study (Additional file 1).

Data analysis

The data were analyzed using the Big Brother software [8]. The following analyses were repeated for the GTV, iGTV, CTV and ICTV data.

Scatterplots were generated in GraphPad Prism (version 5.00, GraphPad Software, San Diego, CA) to present the range of delineated volumes. Using the Big Brother

software we calculated the average volume of the (i)GTV and CI_{gen} for each patient [19]. The CI_{gen} is a measure of overlap of the delineated volumes and is defined as the ratio of the sum over all observer pairs of the volumes common to both observers and the sum over all observer pairs of the encompassing volumes (volume delineated by at least one of the two observers) [19]. CI_{gen} ranges from 0 to 1, where 1 indicates full overlap of the delineated volumes from all observers and 0 indicates no overlap. To assess the accuracy of CI_{gen} we repeated its calculation a number of times equal to the number of observers, leaving out a different observer each repetition. The range of results from this leave-one-out procedure was reported. To test for significant differences in average volumes, we used a two-sided Wilcoxon signed-rank test (32 pairs, significance level $\alpha = 0.05$) using SPSS (version 22.0.0.2, IBM, New York).

To determine the local observer delineation variation per specific area of the (i)GTV or (i)CTV, we calculated for each patient the median surface, i.e. the surface of the volume that was included by at least 50% of the observers [20]. The median surface was sampled with approximately equidistant (0.5 mm) points. For each point on the median surface, the perpendicular distance to each delineated (i)GTV or (i)CTV was measured. When a delineated surface was not within 2 cm perpendicular to a point on the median surface, the closest distance from that delineated surface to the reference point on the median surface was used instead. For each point on the median surface, the local observer variation was calculated, defined as the *SD* of the perpendicular distances at that point (local *SD*). Per patient, the overall observer variation (overall *SD*) was calculated. The overall *SD* was defined as the root-mean-square of the local *SD*s. Similar as for the CI_{gen} , the overall *SD* was repetitively calculated in a leave-one-out procedure and the range was reported.

The answers to the questionnaire were plotted in a scatterplot using GraphPad Prism and the ratings of the difficulty of the delineation between both stages were compared.

Results

Eight observers from six different institutions submitted all GTV and iGTV delineations. The analyses of the delineations reported in this results section were performed on the (i)GTV. The results from the (i)CTV are presented in Additional file 2.

Delineations

Visual inspection of the delineations revealed considerable interobserver variations (Figs. 2.1 and 2.2). The ratio of the largest to the smallest delineated GTV and iGTV was 6.8 and 16.5, respectively, both in patient 3. The iGTV volumes were significantly larger than the GTV volumes by 34% ($p = 0.036$). However, for two observers, the delineated iGTV was smaller than the delineated GTV in all four patients (observers 2 and 5; Fig. 2.3) and for patient 2 two additional observers (6 and 7) also delineated a smaller iGTV than GTV. Observer 7 reported that their iGTV was not based on the ten separate respiratory phases of the 4DCT, due to poor image quality. But as the iGTV was delineated on the Ave-IP of the 4DCT, the iGTV still contained 4DCT information. The CI_{gen} was larger for the GTV (mean $CI_{gen} = 0.37$) than for the iGTV (mean $CI_{gen} = 0.27$) for all four patients, indicating a better overlap of volumes in 3D delineation than in 4D delineation (Table 2.1).

Table 2.1. The average delineated volumes, overall SDs and CI_{gen} for all 4 patients.

Patient		GTV (range*)	iGTV (range*)
1	Average volume, cm ³	36.71 (14.02–75.87)	41.80 (11.85–89.99)
	Overall SD, cm	0.70 (0.47–0.72)	0.71 (0.60–0.72)
	CI_{gen}	0.34 (0.31–0.37)	0.29 (0.26–0.31)
2	Average volume, cm ³	20.26 (7.06–45.21)	20.57 (4.67–67.86)
	Overall SD, cm	0.84 (0.70–0.88)	0.90 (0.37–0.90)
	CI_{gen}	0.22 (0.20–0.27)	0.20 (0.17–0.27)
3	Average volume, cm ³	10.36 (2.91–19.92)	32.38 (5.67–93.58)
	Overall SD, cm	0.48 (0.42–0.51)	0.89 (0.77–0.94)
	CI_{gen}	0.34 (0.30–0.37)	0.16 (0.12–0.19)
4	Average volume, cm ³	52.32 (34.18–76.72)	65.52 (21.48–119.09)
	Overall SD, cm	0.43 (0.38–0.44)	0.68 (0.58–0.70)
	CI_{gen}	0.59 (0.57–0.62)	0.45 (0.42–0.50)
Overall for all patients	Average volume, cm ³	29.91	40.07 [†] ($p = 0.036$)
	Overall SD, cm [‡]	0.63	0.80
	CI_{gen}	0.37	0.27

* Range over eight delineations (average volume) or over results of the leave-one-out analysis (overall SD and CI_{gen}).

† Note that this overall SD was calculated as the root-mean-square of the four overall SDs from the four patients.

‡ Two-sided Wilcoxon signed-rank test

Abbreviations: GTV = gross tumor volume; iGTV = internal gross tumor volume; SD = standard deviation; CI_{gen} = generalized conformity index.

Local observer variation

There was a large local *SD* (local observer variation) at the laterodorsal borders of the GTV and iGTV of patients 1–3 (Figs. 2.1 and 2.4a–b), reflecting the location of the biliary stent. Some observers did, and some did not include the biliary stent in the (i)GTV. The biliary stent was included most often in patient 3: by six observers in the GTV and by four observers in the iGTV (Additional file 3). Especially in patient 2, large local variation was seen. The suspicious pathologically enlarged lymph node in the portocaval space was incorporated in the GTV by five (observers 1–5) and in the iGTV by four (observers 1,3,4 and 8) observers. The suspicious pathologically enlarged lymph node along the common hepatic artery was included in the GTV by three (observers 2–4) and in the iGTV by two (observers 3 and 4) observers (Additional file 3). Also for patient 2, only observer 2 included all the misplaced fiducial markers in the GTV and only observer 1 included all the fiducial markers in the iGTV. For all patients, there was some variation in including the fiducial markers in the delineated volume (Additional file 3). Also, the caudal side of tumors had larger local *SDs* than the other areas of the tumors (Fig. 2.4).

Overall observer variation

The overall observer variation, represented by the overall *SDs* of the (i)GTV, was smaller for the GTV delineations (*SD* = 0.84) compared with the iGTV delineations (*SD* = 0.90) for all four patients (Table 2.1). Due to the observed discrepancy in including the suspicious pathologically enlarged lymph nodes in patient 2, we recalculated the overall *SD* while excluding the portocaval lymph node; the overall *SD* decreased from 0.84 to 0.72 cm for the GTV and from 0.90 to 0.49 cm for the iGTV.

Questionnaire

Seven observers filled out the structured part of the questionnaire; eight observers the open questions. With a mean score of 3.6 for the difficulty of the delineations in both the 3DCT-stage and 4DCT-stage (Fig. 2.5), the observers did not consider the iGTV (using 4DCT) delineation more difficult than the GTV (using 3DCT) delineation. Of the eight observers, one radiation oncologist only just started to treat patients with pancreatic cancer. The remaining seven observers treat on average three to four pancreatic cancer patients per year at their institution (range 1–7.5) and on average they had 5.4 years of experience in delineating pancreatic tumors (range 2–12.5). On average, the observers treated one patient (range 0–4) with pancreatic cancer within the PREOPANC trial. Two observers mentioned in the remarks section that the long

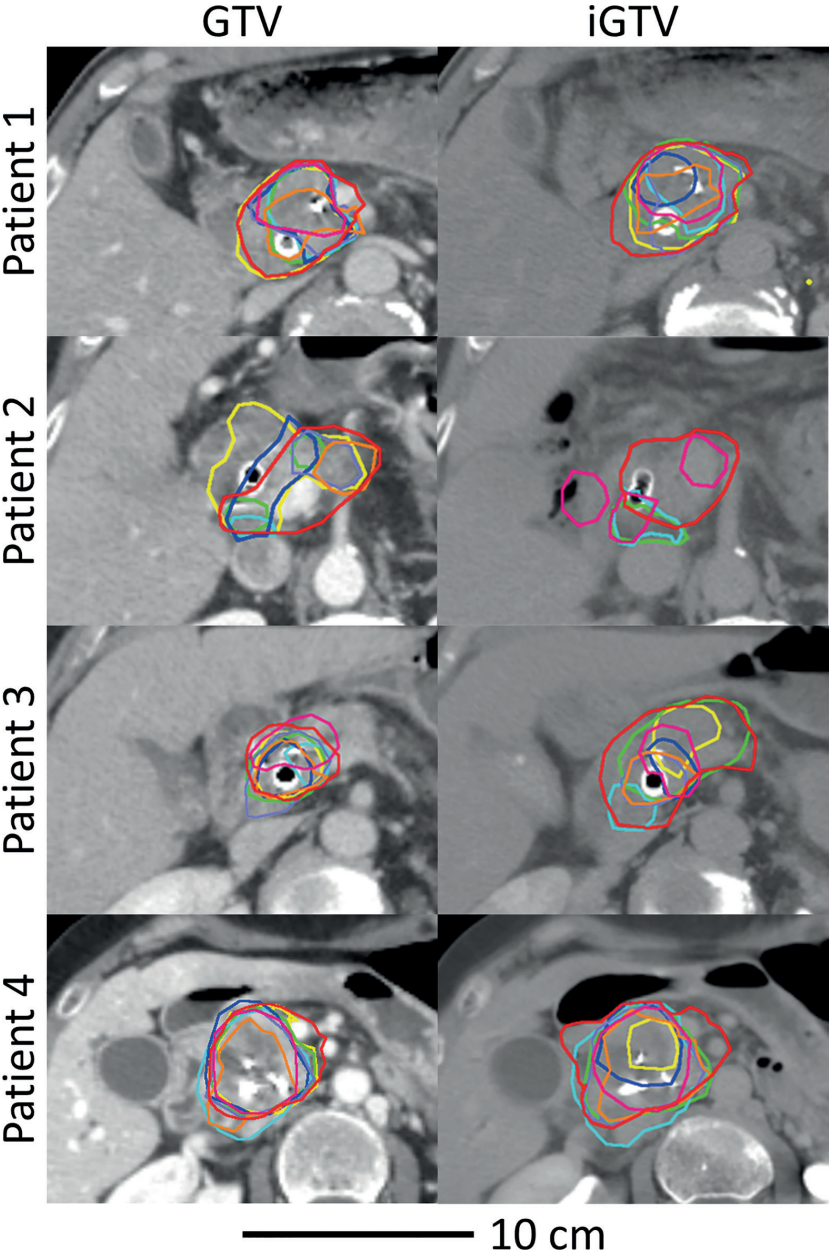


Figure 2.1. For the four patients, delineations of GTV projected onto an axial 3DCT slice (**left**) and iGTV projected onto an axial 4DCT Ave-IP slice (**right**) for all eight observers. Colors are related to observers and are similar for Figs 2.2, 2.3 and 2.5.

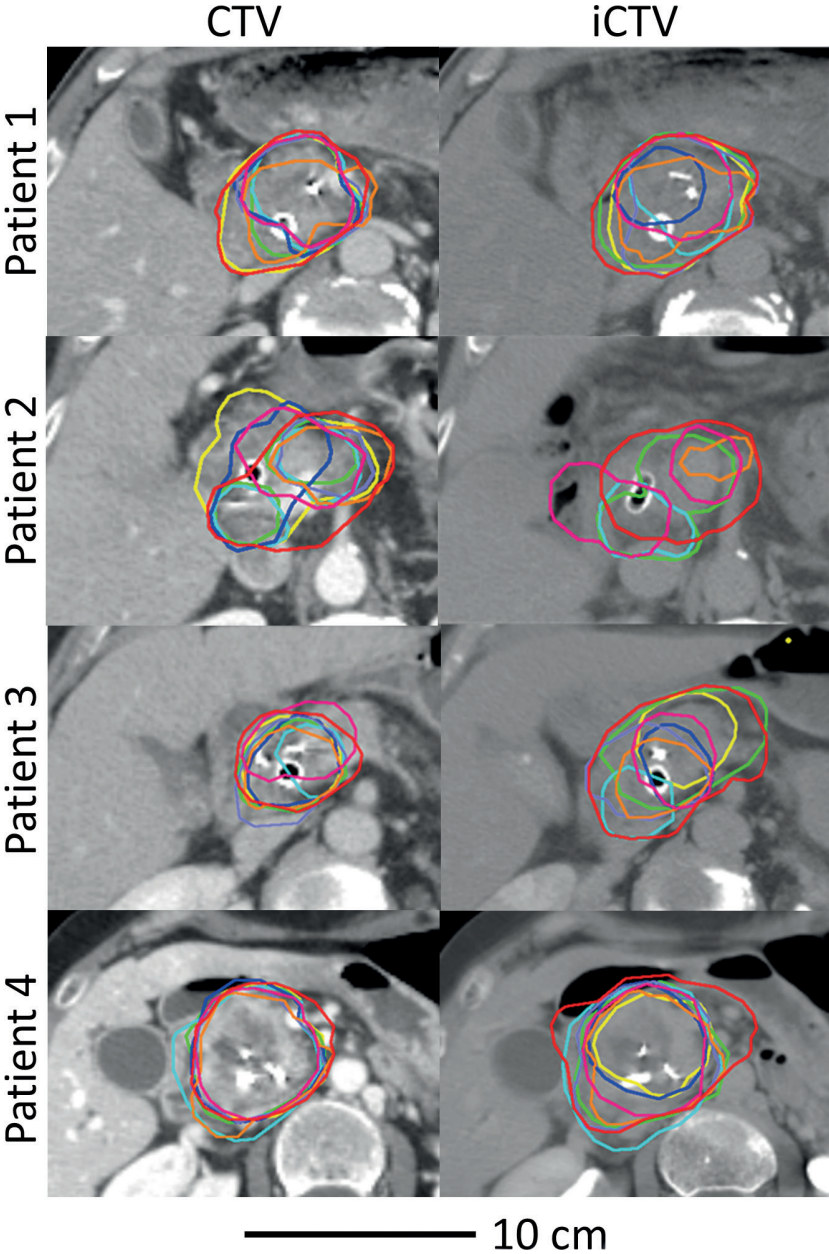


Figure 2.2. For the four patients, the expansion of the CTV projected onto an axial 3DCT slice (**left**) and iCTV projected onto an axial 4DCT Ave-IP slice (**right**) for all eight observers. Colors are related to observers and are similar for Figs 2.1, 2.3 and 2.5.

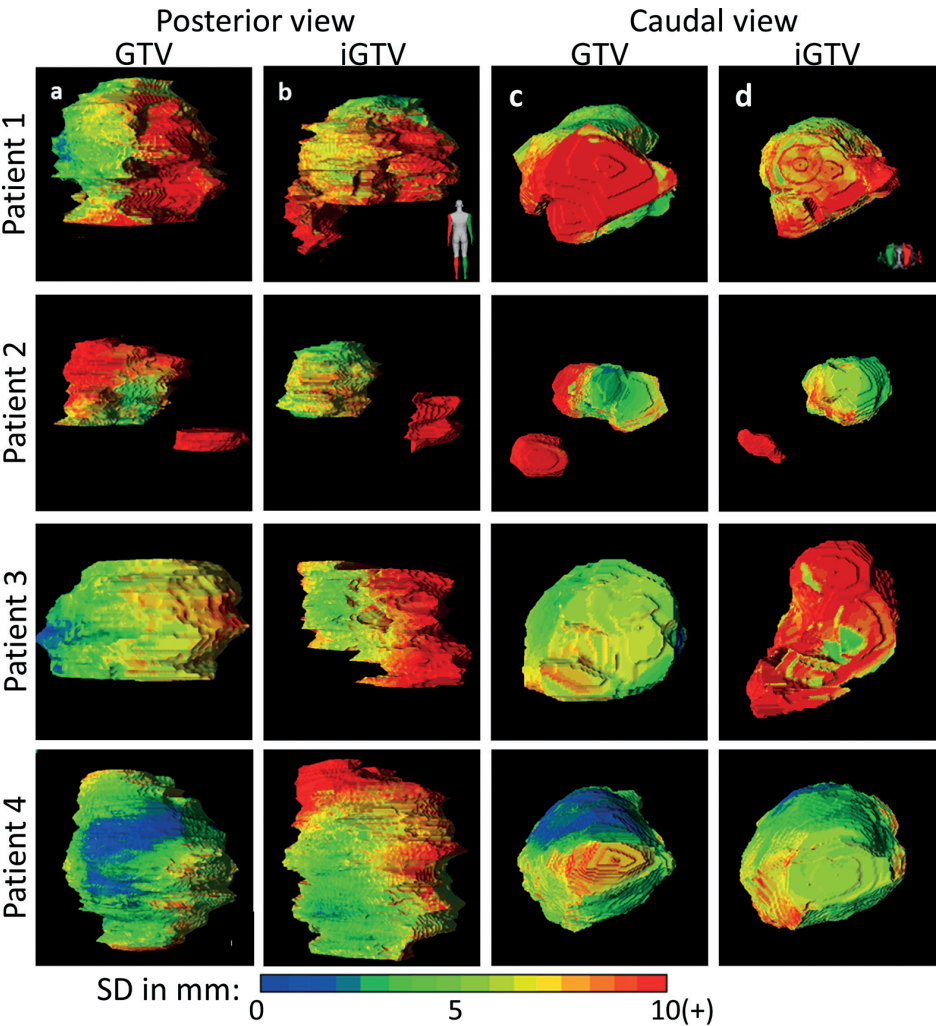


Figure 2.4. For the four patients, the local observer variation in color expressed in local SD (centimeters) of the delineations of the GTV in posterior (a) and caudal view (c) and the iGTV in posterior (b) and caudal view (d) projected onto the median surface. Red indicates a local SD ≥ 0.94 cm.

interval between the diagnostic scan and planning CT scan (average six weeks) made interpretation more challenging.

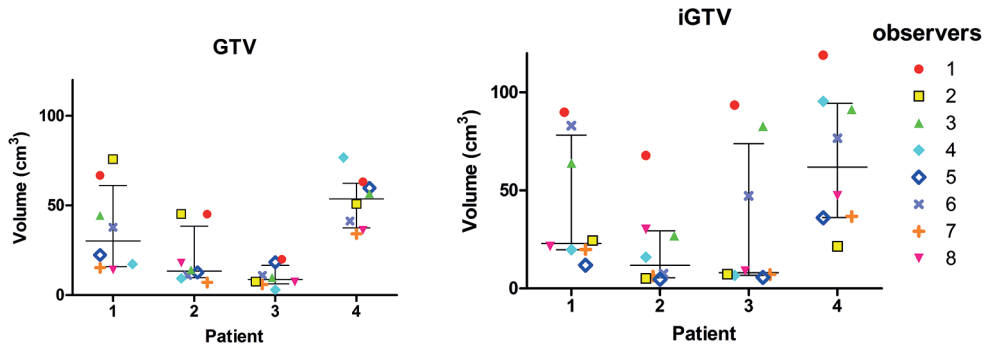


Figure 2.3. Scatterplots of GTV (left) and iGTV (right) of all four patients with the median and 25th and 75th percentile represented by the horizontal lines. Colors are related to observers and are similar for Figs. 2.1, 2.2 and 2.5.

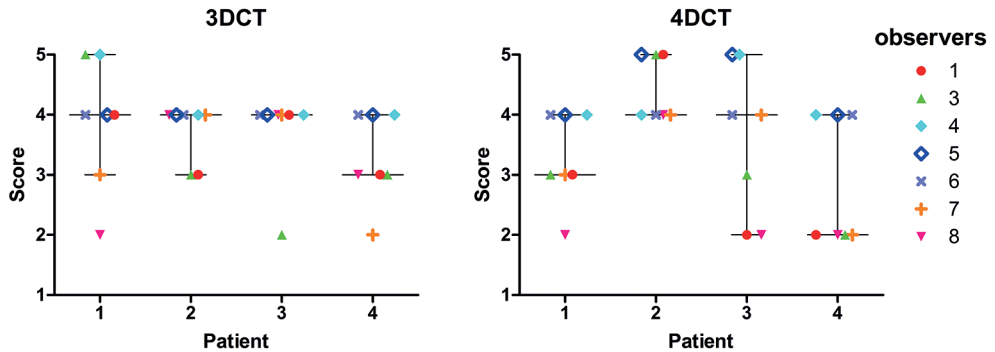


Figure 2.5. Scatterplot of the delineation difficulty rating by seven observer (observer 2 missing), showing the range, the median, 25th, 75th percentile for eight delineations. Rating varied between 1 = very easy; 2 = easy; 3 = moderate; 4 = difficult; 5 = very difficult. Colors are related to observers and are similar for Figs 2.1, 2.2 and 2.3.

Discussion

This multi-institutional delineation study is the first to quantify the interobserver variation on both 3DCT and 4DCT. Also, contrary to earlier studies, this study is performed with more than two patients with (borderline) resectable pancreatic cancer. A considerable variation among observers was observed in both GTV (using 3DCT) and iGTV (using 4DCT) delineations. The ratio of the largest to the smallest delineated volume was far larger for iGTV than for GTV, with significantly larger average volumes for the iGTV. Furthermore, the GTV delineations had larger CI_{gen} and smaller overall SDs in all patients compared to the iGTV. The largest variation in delineation was

seen close to biliary stents and suspicious pathologically enlarged lymph nodes. Previous studies in pancreatic cancer also showed a large interobserver variation on 3DCT with a comparable ratio of largest to smallest GTV of 3–9 [13–15]. The observed interobserver variation is large compared to studies performed in several other organs such as breast, larynx, and lung, which reported a CI_{gen} of larger than 0.6 [19, 21, 22].

The average iGTV volumes were significantly larger than the GTV volumes by 34%. This is similar compared to previous studies in pancreatic cancer, where the iGTV was 25–27.6% larger than the GTV [2, 23]. As the iGTV should incorporate the GTV in all respiratory phases, this result can be expected. However, unexpectedly, in several cases in our study, observers delineated a smaller iGTV than GTV. This may be a result of a large intraobserver variation, which was not assessed in this study. Alternatively, it could be a result of a difference in image quality between the 3DCT and 4DCT images. It is known that inaccuracies in delineation of the tumor may be due to poorly defined tumor edges on CT images [1, 2, 13–15].

The 4DCT delineations had a larger interobserver variation than the 3D delineations, as reflected in the larger overall SDs of the iGTV compared to the GTV in all four patients and the smaller CI_{gen} . This may be the result of poor visibility of the tumor on the various respiratory phases of the 4DCT. Poor visibility can lead to bigger uncertainty and thus larger target volumes and variation in the delineation. The larger interobserver variation on 4DCT may counteract the advantage of the ITV concept: accounting for the respiratory motion. Other delineation approaches with improved contrast between tumor and surrounding tissue to define the tumor borders and including the respiration motion should be investigated such as midventilation and particularly the midposition approach. Previous research showed that a midventilation approach results in significant PTV reduction and significant dose reductions to OARs compared to the iGTV approach, although the delineation process had not been investigated yet [6].

The largest local variation was seen at the laterodorsal side of the (i)GTV, corresponding to the location of the biliary stent. Some observers included the stent in the (i)GTV, whereas others excluded the stent. Also, some observers included the stent only in the GTV but not in the iGTV. In the literature, there is no guideline prescribing to include or exclude the biliary stent in the (i)GTV and none was given in the protocol instructions of the PREOPANC trial. The caudal side of the (i)GTV also showed large variations in delineations, similar to a previous study of Caravatta et al. [13].

Also, large variations in the delineation of the suspicious pathologically enlarged lymph nodes around the tumor were seen. The protocol prescribes to include all neighboring suspicious pathological lymph nodes. The reason for the large local

variation that was found around these lymph nodes could be due to misinterpretation or ambiguity of protocol instructions, or poor compliance with the protocol instructions. To increase interobserver agreement, consensus on the delineation of pancreatic tumors, pathologically enlarged lymph nodes, and biliary stents should be achieved among radiation oncologists. Previous research in other organs showed that national consensus guidelines and a delineation atlas may result in reduction of the interobserver delineation variation [10, 24].

To optimize tumor visibility, the repetition of the diagnostic scan in treatment position after stenting and placement of the fiducial markers may be a step forward. For the patients in our study, registration between the diagnostic CT and the planning CT was not performed because of a different position of the patient and a different anatomy as a result of the placement of the biliary stents and fiducial markers between both scans. Image registration between the diagnostic CT scan and planning CT scan may improve accuracy in target delineation and reduce interobserver variation as seen for other tumor sites [25-27].

It is well known that pancreatic tumors are difficult to distinguish from normal pancreatic tissue on CT scans [2, 28-30]. Therefore, exploitation of other imaging modalities, such as MRI and PET-CT may be a step forward to reduce the variation in delineation of pancreatic tumors. Indeed, other studies have shown that additional imaging, such as MRI and PET-CT, may be helpful in the delineation of pancreatic tumors [31].

Limitations

Delineations were only performed once, and we could not investigate the intraobserver variation. Furthermore, we had a limited number of responding observers, and only a limited number of patients were included. Also, the observers had little experience in the delineation of pancreatic tumors, due to the small number of pancreatic cancer patients eligible for radiotherapy. However, this is typical for many radiation oncologists and hence the found observer variations should be representative for such radiation oncologists. The time interval between diagnostic CT and planning CT scan was on average six weeks and the patients were not scanned in treatment position; therefore, anatomical changes (including placement of the biliary stent) occurred between both scans and scans were not registered. This made it challenging to delineate the (i)GTV. However, this is a typical situation in clinical practice in many hospitals since the diagnostic CT is obtained before histological diagnosis, while therapeutic measures such as stenting are performed after the diagnostic CT scan. Finally, fiducial markers in patient 2 were mistakenly not placed inside the tumor, which may have put some

observers on the wrong track and contributed to the large interobserver variation seen in this patient.

Conclusion

This study showed a considerable interobserver variation in delineation of pancreatic tumors, larger for 4DCT than for 3DCT delineation. The local variation was largest around the biliary stent and suspicious pathologically enlarged lymph nodes. In the future, the addition of other imaging modalities, such as MRI and PET-CT may help decrease observer variation.

References

1. Mancosu P, Bettinardi V, Passoni P et al. Contrast enhanced 4D-CT imaging for target volume definition in pancreatic ductal adenocarcinoma. *Radiother. Oncol.* 2008; 87(3):339–342.
2. Cattaneo GM, Passoni P, Sangalli G et al. Internal target volume defined by contrast-enhanced 4D-CT scan in unresectable pancreatic tumour: Evaluation and reproducibility. *Radiother. Oncol.* 2010; 97(3):525–529.
3. Mori S, Hara R, Yanagi T et al. Four-dimensional measurement of intrafractional respiratory motion of pancreatic tumors using a 256 multi-slice CT scanner. *Radiother. Oncol.* 2009; 92(2):231–237.
4. Lens E, van der Horst A, Kroon PS et al. Differences in respiratory-induced pancreatic tumor motion between 4D treatment planning CT and daily cone beam CT, measured using intratumoral fiducials. *Acta Oncol.* 2014; 53(9):1257–1264.
5. ICRU. Prescribing, recording and reporting photon-beam intensity modulated radiation therapy (IMRT) (ICRU Report 83), 2010. doi:10.1093/jicru/ndq025.
6. Lens E, van der Horst A, Versteijne E et al. Dosimetric advantages of midventilation compared with internal target volume for radiation therapy of pancreatic cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 92(3):675–682.
7. van Mourik AM, Elkhuzen PHM, Minkema D et al. Multiinstitutional study on target volume delineation variation in breast radiotherapy in the presence of guidelines. *Radiother. Oncol.* 2010; 94(3):286–291.
8. Steenbakkers RJHM, Duppen JC, Fitton I et al. Observer variation in target volume delineation of lung cancer related to radiation oncologist-computer interaction: a “Big Brother” evaluation. *Radiother. Oncol.* 2005; 77(2):182–190.
9. Genovesi D, Ausili Cefaro G, Trignani M et al. Interobserver variability of clinical target volume delineation in soft-tissue sarcomas. *Cancer/Radiotherapie* 2014; 18(2):89–96.
10. Nijkamp J, de Haas-Kock DFM, Beukema JC et al. Target volume delineation variation in radiotherapy for early stage rectal cancer in the Netherlands. *Radiother. Oncol.* 2012; 102(1):14–21.
11. Villeirs GM, van Vaerenbergh K, Vakaet L et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther. Onkol.* 2005; 181(7):424–430.
12. Goodman KA, Regine WF, Dawson LA et al. Radiation therapy oncology group consensus panel guidelines for the delineation of the clinical target volume in the postoperative treatment of pancreatic head cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 83(3):901–908.
13. Caravatta L, Macchia G, Mattiucci GC et al. Inter-observer variability of clinical target volume delineation in radiotherapy treatment of pancreatic cancer: A multi-institutional contouring experience. *Radiat. Oncol.* 2014; 9(1):198.
14. Yamazaki H, Nishiyama K, Tanaka E et al. Dummy run for a phase II multi-institute trial of chemoradiotherapy for unresectable pancreatic cancer: Inter-observer variance in contour delineation. *Anticancer Res.* 2007; 27(4 C):2965–2971.
15. Fokas E, Clifford C, Spezi E et al. Comparison of investigator-delineated gross tumor volumes and quality assurance in pancreatic cancer: Analysis of the pretrial benchmark case for the SCALOP trial. *Radiother. Oncol.* 2015; 117(3):432–437.
16. Versteijne E, van Eijck CHJ, Punt CJA et al. Preoperative radiochemotherapy versus immediate surgery for resectable and borderline resectable pancreatic cancer (PREOPANC trial): study protocol for a multicentre randomized controlled trial. *Trials* 2016; 17(1):127.
17. van der Horst A, Wognum S, Dávila Fajardo R et al. Interfractional position variation of pancreatic tumors quantified using intratumoral fiducial markers and daily cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 87(1):202–208.

18. Dávila Fajardo R, Lekkerkerker SJ, van der Horst A et al. EUS-guided fiducial markers placement with a 22-gauge needle for image-guided radiation therapy in pancreatic cancer. *Gastrointest. Endosc.* 2014; 79(5):851–855.
19. Kouwenhoven E, Giezen M, Struikmans H. Measuring the similarity of target volume delineations independent of the number of observers. *Phys. Med. Biol.* 2009; 54(9):2863–2873.
20. Deurloo KEI, Steenbakkers RJHM, Zijp LJ et al. Quantification of shape variation of prostate and seminal vesicles during external beam radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2005; 61(1):228–238.
21. den Hartogh MD, Philippens MEP, van Dam IE et al. MRI and CT imaging for preoperative target volume delineation in breast-conserving therapy. *Radiat. Oncol.* 2014; 9(1):63.
22. Jager EA, Kasperts N, Caldas-Magalhaes J et al. GTV delineation in supraglottic laryngeal carcinoma: interobserver agreement of CT versus CT-MR delineation. *Radiat. Oncol.* 2015; 10:26.
23. Goldstein SD, Ford EC, Duhon M et al. Use of Respiratory-Correlated Four-Dimensional Computed Tomography to Determine Acceptable Treatment Margins for Locally Advanced Pancreatic Adenocarcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2010; 76(2):597–602.
24. Rasch CRN, Steenbakkers RJHM, Fitton I et al. Decreased 3D observer variation with matched CT-MRI, for target delineation in nasopharynx cancer. *Radiat. Oncol.* 2010; 5:21.
25. Yang DS, Yoon WS, Lee JA et al. The effectiveness of gadolinium MRI to improve target delineation for radiotherapy in hepatocellular carcinoma: A comparative study of rigid image registration techniques. *Phys. Medica* 2014; 30(6):676–681.
26. Djan I, Petrović B, Erak M et al. Radiotherapy treatment planning: benefits of CT-MR image registration and fusion in tumor volume delineation. *Vojnosanit. Pregl.* 2013; 70(8):735–739.
27. Bird D, Scarsbrook AF, Sykes J et al. Multimodality imaging with CT, MR and FDG-PET for radiotherapy target volume delineation in oropharyngeal squamous cell carcinoma. *BMC Cancer* 2015; 15(1):844.
28. Fletcher JG, Wiersema MJ, Farrell MA et al. Pancreatic malignancy: value of arterial, pancreatic, and hepatic phase imaging with multi-detector row CT. *Radiology* 2003; 229(1):81–90.
29. McNulty NJ, Francis IR, Platt JF et al. Multi--detector row helical CT of the pancreas: effect of contrast-enhanced multiphasic imaging on enhancement of the pancreas, peripancreatic vasculature, and pancreatic adenocarcinoma. *Radiology* 2001; 220(1):97–102.
30. Li XX, Liu NB, Zhu L et al. Consequences of additional use of contrast-enhanced ¹⁸F-FDG PET/CT in target volume delineation and dose distribution for pancreatic cancer. *Br. J. Radiol.* 2015; 88(1051):20140590.
31. Dalah E, Moraru I, Paulson E et al. Variability of target and normal structure delineation using multimodality imaging for radiation therapy of pancreatic cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):633–640.

Additional file 1

Questionnaire

1. *How did you experience the delineation of the pancreatic tumor of patient 1 on 3DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult
2. *How did you experience the delineation of the pancreatic tumor of patient 2 on 3DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult
3. *How did you experience the delineation of the pancreatic tumor of patient 3 on 3DCT?*
 - d. Very Easy
 - e. Easy
 - f. Moderate
 - g. Difficult
 - h. Very difficult
4. *How did you experience the delineation of the pancreatic tumor of patient 4 on 3DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult

5. *How did you experience the delineation of the pancreatic tumor of patient 5 on 4DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult
6. *How did you experience the delineation of the pancreatic tumor of patient 6 on 4DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult
7. *How did you experience the delineation of the pancreatic tumor of patient 7 on 4DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult
8. *How did you experience the delineation of the pancreatic tumor of patient 8 on 4DCT?*
 - a. Very Easy
 - b. Easy
 - c. Moderate
 - d. Difficult
 - e. Very difficult
9. *How many years' experience do you have in delineation pancreatic tumors for radiotherapy?*
10. *How many patients with pancreatic tumors do you treat per year?*
11. *Have you treated patients within the PREOPANC study?*

Additional file 2

Data of CTV and iCTV

Table 2.A. The volumes, overall standard deviations and conformity indexes of all 4 patients.

Patient		CTV (range*)	iCTV (range*)
1	Average volume (cm ³)	79.21 (36.04–145.96)	89.15 (31.70–165.48)
	Overall SD (cm)	0.75 (0.60–0.77)	0.82 (0.69–0.85)
	CI_{gen}	0.43 (0.41–0.46)	0.38 (0.35–0.41)
2	Average volume (cm ³)	52.99 (24.97–106.90)	53.05 (16.59–131.79)
	Overall SD (cm)	0.90 (0.80–0.95)	1.05 (0.39–1.05)
	CI_{gen}	0.32 (0.29–0.37)	0.30 (0.27–0.34)
3	Average volume (cm ³)	28.66 (11.85–46.76)	68.26 (19.36–169.21)
	Overall SD (cm)	0.52 (0.45–0.55)	0.99 (0.88–1.04)
	CI_{gen}	0.46 (0.43–0.50)	0.23 (0.19–0.26)
4	Average volume (cm ³)	103.86 (75.13–147.46)	126.75 (49.2–204.04)
	Overall SD (cm)	0.44 (0.39–0.46)	0.72 (0.62–0.74)
	CI_{gen}	0.65 (0.63–0.67)	0.50 (0.47–0.54)
Overall for all patients	Average volume (cm ³)*	66.18	84.30 [‡] (p = 0.045)
	Overall SD (cm) [‡]	0.68	0.91
	CI_{gen}	0.46	0.35

* Range over 8 delineation (average volume) or results of leave-one-out analysis (overall SD and CI_{gen}).

† Note that the overall SD was calculated as the root-mean-square of the four SDs from the four patients

‡ Two-sided Wilcoxon signed-rank test

Abbreviations: CTV = clinical target volume; iCTV = internal clinical target volume; SD = standard deviation;

CI_{gen} = generalized conformity index.

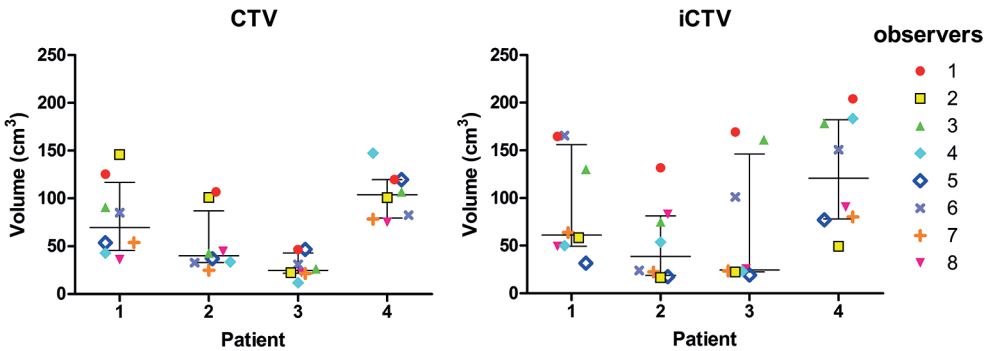


Figure 2.A. Scatterplots of CTV (left) and iCTV (right) of all four patients with the median, 25th and 75th percentile represented by the horizontal lines. Colors are related to observers and are similar for Figs. 2.1, 2.2, 2.3 and 2.5 in this paper.

Additional file 3

Suspicious pathological lymph nodes, stents, and fiducials

Table 2.B. Number of observers who included the suspicious pathological lymph nodes in the (i)GTV according the diagnostic CT report.

Patient	Location suspicious pathological lymph node	GTV	iGTV
2	Portocaval	5	4
2	Along common hepatic artery	3	2
4	Along tumor, not characterized	2	2

Table 2.C. Number of observers who included the biliary stent or percutaneous biliary drainage in the (i)GTV at least 50% of the stent in at least three slices.

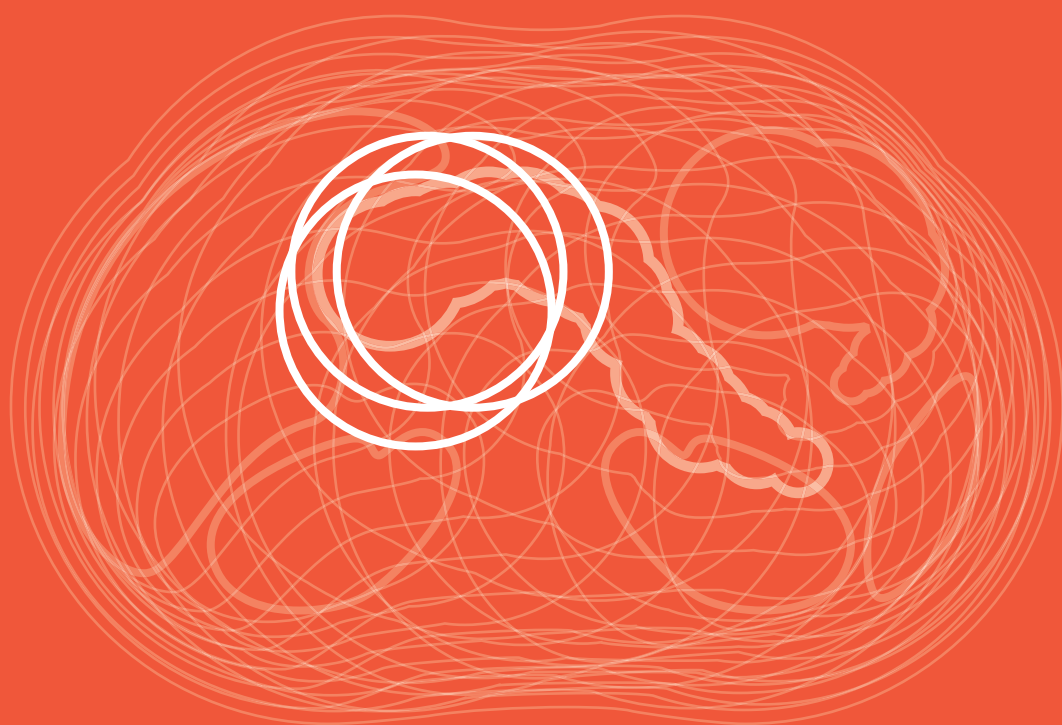
Patient	GTV-CT	iGTV-CT
1	3	5
2	2	1
3	6	4
4	8	7

Table 2.D. The number of (i)GTV including the fiducial over the number of fiducials multiplied by number of observers ($3 \times 8 = 24$). Partly included fiducials were counted as delineated.

Patient	GTV-CT	iGTV-CT
1	24/24 (100%)	23/24 (100%)
2*	5/24 (21%)	4/24 (17%)
3	22/24 (92%)	16/24 (67%)
4	24/24 (100%)	21/24 (88%)

* In patient 2, the fiducials were mistakenly not implanted in the tumor.





CHAPTER 3

Delineation on CT+MRI

Addition of MRI improves interobserver variation in CT-based pancreatic tumor delineation

Oliver J Gurney-Champion
Eva Versteijne
Astrid van der Horst
Eelco Lens
Heidi Rütten
Hanne D Heerkens
Gabriel MRM Paardekooper
Maaïke Berbee
Coen RN Rasch
Jaap Stoker
Marc RW Engelbrecht
Marcel van Herk
Aart J Nederveen
Remy Klaassen
Hanneke WM van Laarhoven
Geertjan van Tienhoven
Arjan Bel

submitted

Abstract

Purpose: To assess whether the addition of magnetic resonance imaging (MRI) alongside the planning computed tomography (CT) scan decreases interobserver variation for target volume delineation in pancreatic cancer patients.

Methods and Materials: Eight observers (radiation oncologists) from six institutions delineated the gross tumor volume (GTV) on 3DCT, and internal GTV (iGTV) on 4DCT of four pancreatic cancer patients, while MRI was available in a second window (CT+MRI). Variations in volume, generalized conformity index (CI_{gen}), and overall observer variation, expressed as standard deviation (SD) of the distances between delineated surfaces, were analyzed. CI_{gen} is a measure of overlap of the delineated (i)GTVs (1 = full overlap, 0 = no overlap). A Wilcoxon signed-rank test (significance level $\alpha = 0.05$) was used to compare these results to those from an earlier study that assessed the interobserver variation by the same observers on the same patients on CT without MRI (CT-only).

Results: The maximum ratios between delineated volumes within a patient were 6.1 and 22.4 for the GTV (using 3DCT) and iGTV (using 4DCT), respectively. The average (root-mean-square) overall observer variations were $SD = 0.41$ cm (GTV) and $SD = 0.73$ cm (iGTV). The mean CI_{gen} was 0.36 for GTV and 0.37 for iGTV. When compared to the (i)GTV delineated on CT-only, the mean volumes of the (i)GTV on CT+MRI were significantly smaller (32%, $p < 0.0005$). The median volumes of the (i)GTV on CT+MRI were included for 97% and 92% in the median volumes of the GTV and iGTV on CT, respectively. Furthermore, CT+MRI showed smaller overall observer variations (GTV and iGTV root-mean-square $SD = 0.59$ cm) in six out of eight delineated structures compared to CT-only (root-mean-square $SD = 0.72$ cm). Finally, although smaller volumes were delineated on CT+MRI, the CI_{gen} was similar in both studies.

Conclusion: The availability of MRI images during target delineation reduced the interobserver variation for 3DCT and 4DCT delineation of target volumes in pancreatic cancer.

Introduction

Radiotherapy for pancreatic cancer aims at delivering a high radiation dose to the tumor while minimizing the dose delivered to the organs at risk (OARs). Several developments have improved the accuracy of radiotherapy of pancreatic cancer. The introduction of intensity modulated radiotherapy and volumetric modulated arc therapy enabled steep dose gradients close to the tumor, reducing the dose to OARs [1, 2]. Also, patient alignment has greatly improved with the introduction of intratumoral fiducial markers combined with daily cone-beam computed tomography (CBCT) [3, 4]. Furthermore, motion management has been improved with the introduction of 4D computed tomography (CT) [5] and breath-holding [6]. One of the remaining major challenges in radiotherapy for pancreatic cancer patients is precise and accurate target volume definition [7-9].

So far, studies on delineation accuracy investigated delineations on 3DCT in 1–2 patients with 11–25 observers. However, no quantitative information was reported on the conformity of the delineated volumes, or on local and overall observer variations expressed as standard deviations (*SD*). Such parameters are especially relevant for determining appropriate planning treatment volume margins [10] and quantifying improvement in interobserver variation after intervention [11, 12]. In an earlier study with eight observers and four patients, we confirmed a substantial interobserver variation when the tumor was delineated on 3DCT and 4DCT based on these parameters [13].

One of the reasons for this considerable interobserver variation may be the poor contrast between pancreatic tumors and the surrounding tissue on CT images. Potentially, MRI provides better tumor visibility [14–18]. For several other organs, MRI has been shown to improve interobserver variation [19–22]. For pancreatic tumors, delineations based on MRI were only studied in single institute studies (1–3 observers), which did not quantify the interobserver variation [23, 24]. To our knowledge, no multi-center study that assesses the value of MRI for target volume delineation in pancreatic cancer patients is available.

The aim of this exploratory study was to evaluate whether there is an added value in offering MRI alongside the planning CT for delineation of the target volume in pancreatic cancer patients. In this study, we quantify the interobserver variation and compare it to our previously published results [13] on CT-based delineation with the same eight observers and four patients.

Materials and Methods

In our previous study (CT-only) [13], eight observers (radiation oncologists) from six Dutch institutions participating in the PREOPANC trial [25] delineated target volumes of four pancreatic cancer patients using only diagnostic and planning CTs. In the current study (CT+MRI), we asked the same observers to repeat the delineations for the same patients, now offering a diagnostic MRI alongside the CTs. All CT-only data presented in this paper as comparison come from the earlier study [13].

Patients

The same patients were selected as in the CT-only study [13]. In that study, we selected the first four patients with histologically proven (borderline) resectable pancreatic ductal adenocarcinoma who participated in the PREOPANC (EudraCT number 2012-003181-40) trial [25] (radiochemotherapy arm) and MIPA (NCT01989000) study. The patients gave written informed consent to both studies, which were approved by the local medical ethics committees (PREOPANC: Erasmus Medical Center, Rotterdam; MIPA: Academic Medical Center, Amsterdam).

Imaging

Patients underwent a diagnostic CT, MRI and planning CT examination. The diagnostic CT scans (contrast-enhanced; CE) were acquired as part of standard patient care at the referring hospitals. Experienced abdominal radiologists from our institution reviewed these scans and considered them adequate for diagnostic purposes.

After diagnosis, all patients received three markers (intratumoral golden fiducial markers) that were visible on CT, but not on MRI [26], as part of our standard treatment protocol [3, 27]. Furthermore, patients 1–3 received metallic biliary stents, and patient 4 received percutaneous biliary drainage. All were placed after the diagnostic CT, but before the MRI and planning CT.

MRI was performed on an Ingenia 3 T scanner (Philips Healthcare, Best, the Netherlands) as part of the MIPA study. Four MRI scans were obtained using various sequences: T1-weighted spoiled gradient echo (T1W GE), CE T1W GE, T2-weighted turbo spin echo (T2W TSE) and diffusion-weighted imaging (DWI) for which the apparent diffusion coefficient map was displayed [28, 29] (examples in Fig. 3.1 e-h). The diagnostic CT and MRI were not registered to the planning CT.

The planning CT scans were acquired at our radiation oncology department on a GE LightSpeed RT 16 scanner (General Electric Company, Waukesha, WI). Two planning CT scans were acquired, a CE 3DCT and a 4DCT. Several image sets were

reconstructed from the 4DCT scan: the ten respiratory phases, average intensity projection (Ave-IP) and maximum intensity projection (MIP). The diagnostic CT and MRI examinations were performed before the start of radiochemotherapy, whereas the planning CT examination was on average eight days (range 6–10 days) after the patients received their first administration of gemcitabine. Further details of all scans are discussed in the Supplementary Materials.

Delineation

Observers delineated the gross tumor volume (GTV) on the 3DCT (CE) and the internal GTV (iGTV) on the Ave-IP reconstruction of the 4DCT. The GTV was defined as the macroscopically visible tumor and pathological lymph nodes. The iGTV was defined as the GTV delineated on the Ave-IP reconstruction, extended to encompass the GTV on the ten respiratory phases of the 4DCT. The current study (CT+MRI), consisted of a 3DCT+MRI and 4DCT+MRI stage, in which the GTV and iGTV, respectively, were delineated under the guidance of MRI. The results were compared to the results from the 3DCT-only and 4DCT-only stages of the CT-only study [13]. For both studies observers received the same instructions on what to delineate, taken from the PREOPANC trial [25].

In both studies, observers received the Big Brother software [30]. The software showed a primary window, in which the (i)GTV was delineated, and a secondary window that could display selected other available images (Table 3.1). When viewed in the secondary window, the displayed slice from the 3DCT and 4DCT image and cursor position, were linked to the primary window (i.e. same slice and a dot indicating cursor position). Furthermore, observers had access to the radiology reports of the diagnostic CT and, in the CT+MRI study, the radiology reports of the MRI. These reports, from experienced abdominal radiologists, described the tumor extent. For the CT report, the associated pathological lymph nodes were also described: two suspicious locoregional lymph nodes in patient 2; “some” (cited) enlarged lymph nodes in patient 4, which were not characterized further.

First, the software for the CT-only study was sent to all observers. Four weeks after an observer returned their CT-only delineations, that observer received a PowerPoint document. In this document, the visibility of pancreatic tumors on the abovementioned MRI scans was discussed for nine pancreatic cancer patients (different from those included in this study). Observers received the software for the CT+MRI study at least six weeks after returning their CT-only delineations.

Table 3.1. Overview of the available images.

	3DCT+MRI	3DCT-only*	4DCT+MRI	4DCT-only*
3DCT	X	X	+	+
4DCT Ave-IP			X	X
4DCT MIP			+	+
4DCT ten phases			+	+
Diagnostic CTs	+	+	+	+
T1W GE	+		+	
CE T1W GE	+		+	
T2W TSE	+		+	
DWI†	+		+	

* These stages are from the CT-only study [13].

† The apparent diffusion coefficient map of the DWI images was shown.

X = images in main window, used for delineation; + = available in secondary window.

Abbreviations: Ave-IP = average intensity projection; MIP = maximum intensity projection;

T1W GE = T1-weighted gradient echo; CE = contrast enhanced; T2W TSE = T2-weighted turbo spin echo;

DWI = diffusion-weighted imaging.

CT+MRI

The analyses were performed using the Big Brother software [30]. A resident radiation oncologist (EV) visually assessed the individually delineated (i)GTVs. She counted the number of observers who included the stents/drain, markers and pathological lymph nodes in the delineated volume. The average volume of the (i)GTVs, and generalized conformity index (CI_{gen}) [31] per patient was calculated with the Big Brother software. CI_{gen} is the sum over all observer pairs of their common volume (delineated by both observers), divided by the sum over all observer pairs of their encompassing volumes (delineated by at least one observer). CI_{gen} is a measure of overlap of the (i)GTVs (1 = full overlap, 0 = no overlap). A median surface was defined using Big Brother. This surface was the 3D closed surface comprising the volume that was included in the (i)GTV by at least 50% of the observers. The median surface was sampled with approximately equidistant (0.5 mm) points in Big Brother. For each point, the perpendicular distances from the median surface to the surfaces of the eight individually delineated (i)GTVs were measured. If the surface of a delineated (i)GTV was not within 2 cm, the distance to the closest point on that surface was used instead. The SD over these eight distances was used as a measure of local observer variation (local SD). The overall observer variation (overall SD) was defined as the root-mean-square of all local SDs. Average overall SDs were calculated as root-mean-square over all patients. To assess

the dependence of the CI_{gen} and overall SD on individual observers, we reported the range of values in a leave-one-out analysis. In this approach, the analyses were repeated eight times, successively leaving out one of the observers.

Comparison CT+MRI with CT-only

The results were compared to results from the CT-only study [13]. All statistical analyses were performed using SPSS (version 22.0.0.2, IBM, New York). Plots were made with GraphPad Prism (version 5.00, GraphPad Software, San Diego, CA). To test for differences in mean delineated target volumes, CI_{gen} and mean overall SD we used a two-sided Wilcoxon signed-rank test between the CT-only and CT+MRI studies (significance level $\alpha = 0.05$). Histograms of the local SD s per patient were compared between the CT-only study and CT+MRI study. Finally, we compared the 3D median surfaces from both studies and calculated for each median surface from the CT-only study the percent of the volume surrounded by the median surface of the CT+MRI study and vice versa.

Results

Observers had on average 10.6 years (range 3–17 years) of experience in radiotherapy. Observer 4 (12 years of experience in radiotherapy) only just started treating pancreatic cancer. The remaining seven had an average of 5.4 years (range 2–12.5 years) of experience treating pancreatic cancer. These observers saw an average of 3.5 pancreatic cancer patients (range 1–7.5 patients) per year. All of the observers had experience with MRI in radiotherapy, and four had experience with MRI for pancreatic cancer.

CT+MRI

The maximum ratio between delineated target volumes within one patient was 6.1 for 3DCT+MRI and 22.4 for 4DCT+MRI (Figs. 3.1 a and c). The mean CI_{gen} was 0.36 for 3DCT+MRI and 0.37 for 4DCT+MRI (Table 3.2). The mean overall observer variation expressed by overall SD was 0.41 cm for 3DCT+MRI and 0.73 cm for 4DCT+MRI (root-mean-square over all patients). Despite the instructions that pathological lymph nodes should be included in the (i)GTV, the different lymph nodes were only delineated by 1–4 (range) out of eight observers (supplemental Table 3.B). Furthermore, the (i)GTVs

showed large variations close to stents/drain, in particular in patient 3 in whom four out of eight observers included the stent for 4DCT+MRI and five out of eight for 4DCT+MRI (supplemental Table 3.C).

Table 3.2. Overview of the average volume, CI_{gen} and overall observer variation (overall SD).

	3DCT+MRI	3DCT-only*	4DCT+MRI	4DCT-only *
Average volume (cm ³)				
Patient 1	13.3 (6.1–32.5)	36.7 (14.0–75.9)	19.3 (10.5–52.5)	41.8 (11.9–90.0)
Patient 2	9.1 (4.5–13.8)	20.3 (7.1–45.2)	18.5 (3.0–51.3)	20.6 (4.7–67.9)
Patient 3	5.7 (1.9–11.8)	10.4 (2.9–19.9)	11.0 (3.2–26.9)	32.4 (5.7–93.6)
Patient 4	53.4 (34.8–96.0)	52.3 (34.2–76.7)	59.8 (38.2–76.9)	65.5 (21.5–119.1)
Mean:	20.4	29.9	27.13	40.07
CI_{gen}				
Patient 1	0.31 (0.29–0.39)	0.34 (0.31–0.37)	0.25 (0.23–0.29)	0.29 (0.26–0.31)
Patient 2	0.34 (0.32–0.42)	0.22 (0.20–0.27)	0.17 (0.15–0.21)	0.20 (0.17–0.27)
Patient 3	0.23 (0.20–0.36)	0.34 (0.30–0.37)	0.20 (0.17–0.26)	0.16 (0.12–0.19)
Patient 4	0.56 (0.53–0.61)	0.59 (0.57–0.62)	0.61 (0.59–0.62)	0.45 (0.42–0.50)
Mean	0.36	0.37	0.31	0.27
Overall SD (cm)				
Patient 1	0.43 (0.36–0.44)	0.70 (0.47–0.72)	0.57 (0.53–0.60)	0.71 (0.60–0.72)
Patient 2	0.40 (0.28–0.43)	0.84 (0.70–0.88)	1.11 (0.87–1.14)	0.90 (0.37–0.90)
Patient 3	0.40 (0.35–0.40)	0.48 (0.42–0.51)	0.62 (0.50–0.66)	0.89 (0.77–0.94)
Patient 4	0.43 (0.37–0.45)	0.43 (0.38–0.44)	0.39 (0.35–0.40)	0.68 (0.58–0.70)
RMS:	0.41	0.63	0.73	0.80

* These results are from the CT-only study (13).

Abbreviations: SD = standard deviation; CI_{gen} = generalized conformity index; RMS = root-mean-square. Between brackets, range of volumes (Average volume) and range of leave-one-out analyses (overall SD and CI_{gen}).

Comparison CT+MRI with CT-only

Delineated (i)GTVs were 32% smaller when based on CT+MRI compared to CT-only ($p < 0.0005$, $Z = -3.826$, Fig. 3.1 a-d, Fig. 3.2 and Table 3.2). There was no significant difference in mean CI_{gen} between the CT+MRI study ($CI_{gen} = 0.34$) and CT-only study ($CI_{gen} = 0.32$; $p = 0.844$, $Z = -0.280$). The mean overall observer variation was not significantly smaller in the CT+MRI study (root-mean-square over patients and (i)GTV overall SD = 0.59 cm) than in the CT-only study (0.72 cm; $p = 0.078$, $Z = -1.820$).

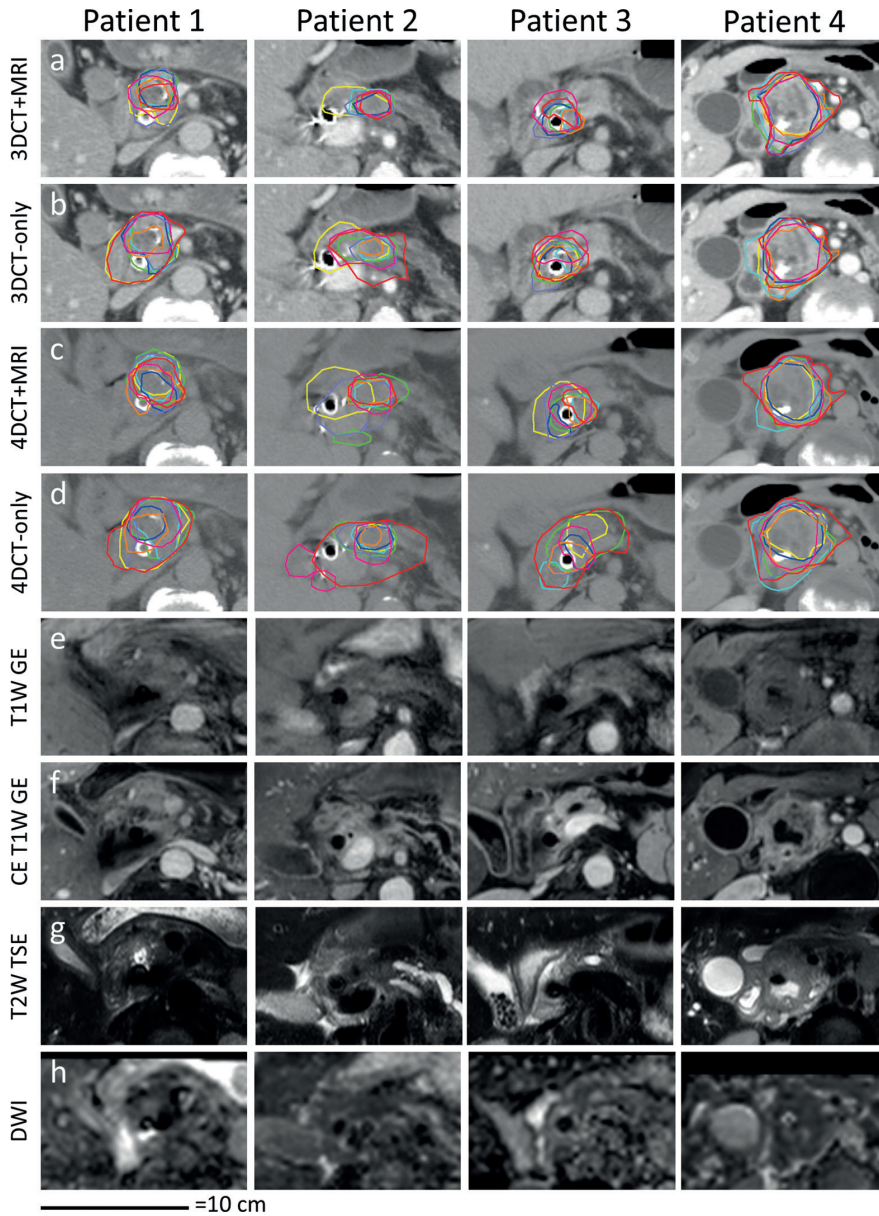


Figure 3.1. Example slices of GTV delineated on 3DCT+MRI (a) or 3DCT-only [13] (b) and iGTV delineated on the Ave-IP using 4DCT+MRI (c) or 4DCT-only [13] (d). The MRI images (e-h) show the tumor in the manually selected corresponding slice. The apparent diffusion coefficient map from the DWI acquisition is shown (h). Abbreviations: T1W = T1-weighted; GE = gradient echo; CE = contrast enhanced; T2W = T2-weighted; TSE = turbo spin echo; DWI = diffusion-weighted imaging.

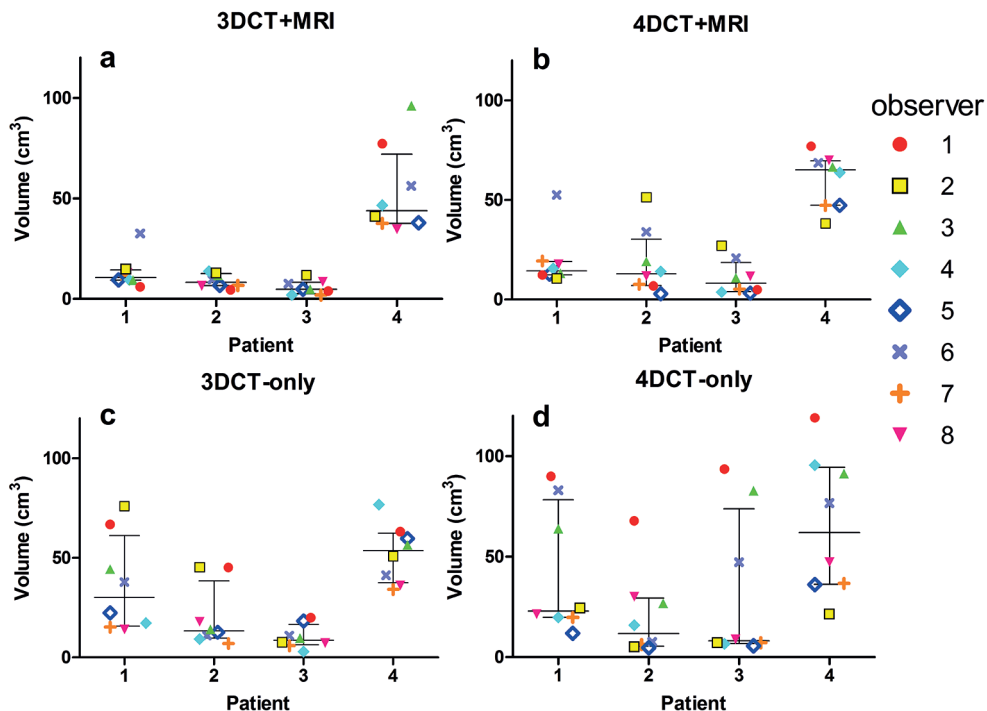


Figure 3.2. Scatter plots of the volume of the delineated (i)GTV on 3DCT+MRI (a), 4DCT+MRI (b), 3DCT-only (c), and 4DCT-only (d) per patient per observer. Median and interquartile ranges are indicated. Observer colors correspond to Fig. 3.1 a-d. Graphs (c) and (d) are from Versteijne et al. [13].

However, they were smaller in six out of eight delineated structures (three out of four GTVs, three out of four iGTVs).

Making MRI available decreased the local observer variation particularly in regions that had a large (> 1 cm) local *SD* in the CT-only study (Fig. 3.3). The histograms of local *SD* reflect this effect (supplemental Fig. 3.A). These histograms show shorter tails at high local *SD* for CT+MRI compared to CT-only, especially for the GTV of patients 1 and 2, and the iGTV of patients 3 and 4.

The median delineated surfaces of the (i)GTV of patient 2 each consisted of two separate volumes (except during the 3DCT+MRI stage), one encompassing the main tumor and one encompassing a suspicious portocaval lymph node (Fig. 3.3). As only 4–5 observers delineated this lymph node (supplemental Table 3.B), the local observer variation was large for that part of the median surface (Fig. 3.3, patient 2: local *SD* > 1 cm). Excluding this lymph node from the median surface resulted in an overall *SD* = 0.72 cm (leave-one-out range: 0.50–0.75 cm) in the 3DCT-only stage,

0.86 cm (0.44–0.91 cm) in the 4DCT+MRI stage and 0.49 cm (0.37–0.51 cm) in the 4DCT-only stage.

The median surfaces from the CT+MRI study were on average for 97% and 92% included within the median surfaces of the CT-only study for the GTV (using 3DCT) and iGTV (using 4DCT), respectively (Fig. 3.4). Vice versa, the median surfaces from the CT-only study were for 56% and 64% included within the median surface of the CT+MRI study.

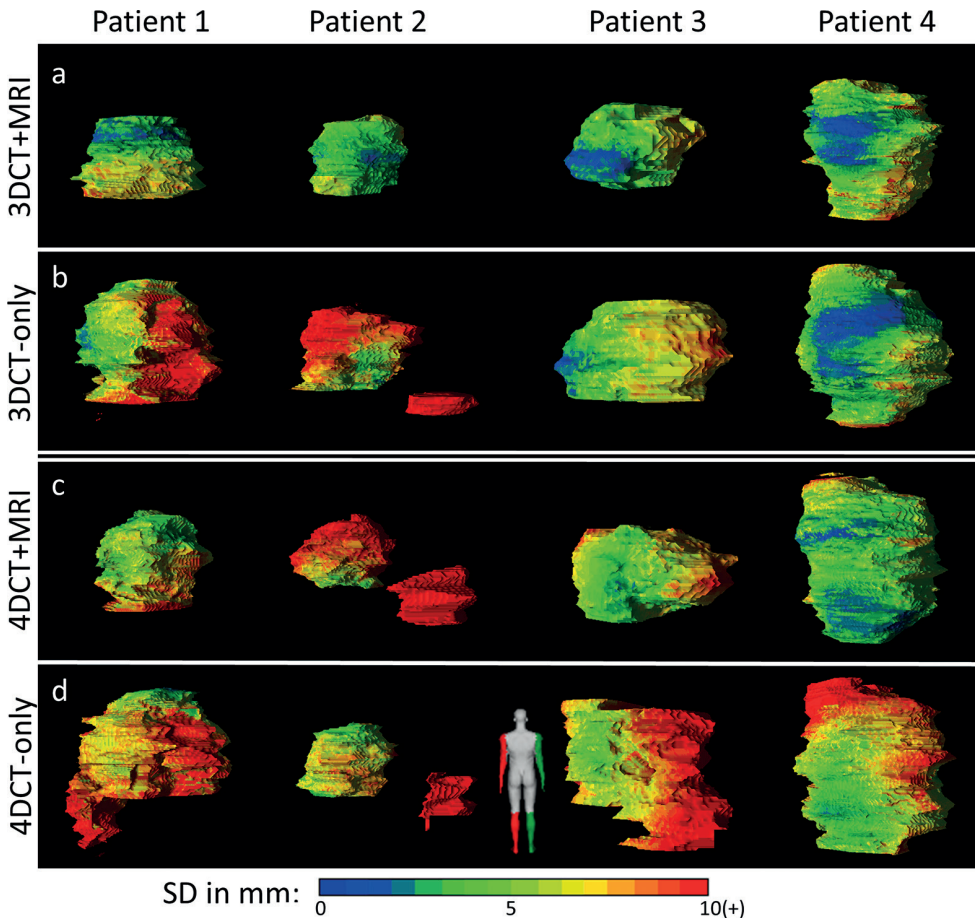


Figure 3.3. The local observer variation (local SD) projected on the median surfaces of the (i)GTV for 3DCT+MRI (a), 3DCT-only (b), 4DCT+MRI (c) and 4DCT-only (d). Colors correlate to the local SD, with red indicating local SD ≥ 9.4 mm. Volumes are viewed posterior of the patient (dummy). Note that for patient 2 the portocaval lymph node was included in (b-d). Figures (b) and (d) are from Versteijne et al [13].

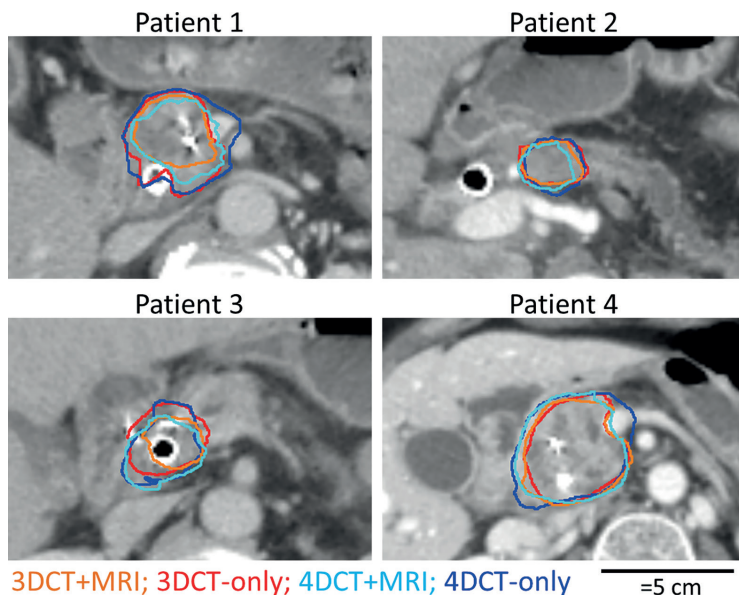


Figure 3.4. For each patient, median surfaces for CT+MRI study (orange and light blue) and the CT-only study (red and dark blue) are shown.

The number of observers who delineated a specific lymph node only varied by a maximum of one observer between the two studies (supplemental Table 3.B). An exception was the portocaval lymph node in patient 2, which was delineated by two observers in the 3DCT+MRI stage, instead of 4–5 observers in all other stages. There was a poor agreement in all stages on whether to include stents/drains in the (i)GTV (supplemental Table 3.C). There was a small difference ($< 17\%$ for all patients) between the CT+MRI study and CT-only study in the percentage of markers included in the (i)GTV (supplemental Table 3.D), except in patient 3. In patient 3 the markers were less often included for CT+MRI than for CT-only delineations.

Discussion

We are the first to show the benefits of offering MRI alongside the planning CT for delineation of pancreatic tumors in a multi-center setting. When MRI was available, the precision, represented by the overall SD , improved in six out of eight delineated structures. Furthermore, the volumes of the (i)GTV decreased significantly. The overlap of delineated (i)GTVs, represented by CI_{gen} , remained similar, despite the

smaller volumes. These findings suggest an extra value of adding MRI for pancreatic tumor delineation.

In our study, as well as in other studies [23, 24], delineated target volumes in pancreatic cancer patients were smaller when (partially) based on MRI than when based on CT only. Similar findings have been reported for other types of cancer [19-21]. There are two possible causes for this decrease in volume. The tumor volume size could be overestimated on CT due to poor tumor contrast and uncertain tumor boundaries. Furthermore, tumor size could be underestimated on (CT+)MRI. Therefore, appropriate clinical target volume (CTV) margins should be determined.

We assessed the number of interactions done in Big Brother with the various MRI sequences and found that observers focused on the CE T1W GE (results not shown). In one study [32], pancreatic tumor sizes were underestimated when assessed on CE MRI. In a different study [23], it was shown that delineations of pancreatic cancer tumors on MRI were larger when based on DWI than when based on CE T1W GE images. More extensive use of the other MRI sequences may improve delineations [18]. Future protocol instructions for MRI-based tumor delineation should be developed taking into account these issues.

The addition of MRI mainly decreased the local *SD* in regions that had large local observer variation (local *SD* > 1cm). The fact that less decrease was seen in other regions of low local observer variation could be a result of the MRI images not being registered to the planning CT images. Translating the MRI information to the CT images may be challenging for the observers. Potentially, registration of the CT and MRI images decreases the overall observer variation further, as shown for brain tumors [33]. As the pancreas and other abdominal organs had deformed between the three image sessions, a deformable registration would be required. The use of deformable image registration for this purpose is not widely validated. Therefore, we chose not to register the images. In the future, MRI and CT scans should be acquired in the treatment position, preferably with markers that are visible on both MRI and CT to guide registration [26].

Clear instructions and consensus on what to delineate decreases interobserver variation, as was shown in various organs [12, 11]. In our study, no specific instructions on whether to include stents/drains into the (i)GTV were given. The large variation on including stents suggests the necessity of such instructions. Our delineation instructions did state that pathological lymph nodes should be included in the (i)GTVs. The large variation concerning the inclusion of lymph nodes in the (i)GTV suggests that observer compliance is also important. As these confusions affected both the CT-only and CT+MRI studies (Tables 3.B-3.D), we believe they did not influence the comparison between the CT-only and CT+MRI delineations.

Several limitations were associated with this paper. Due to the relatively large amount of pancreatic tumor delineations done (four patients, four times) in a relatively short time, observers may have improved their skills between both studies. Furthermore, despite the gap of at least six weeks between both studies, observers may have recognized patients during the CT+MRI study. Both factors could result in smaller observer variation for CT+MRI. Furthermore, the MRI was not obtained as part of the radiotherapy treatment. When MRI is obtained for radiation treatment, different settings or sequences may be preferred with higher resolutions, such as e.g. alternating repetition time balanced steady-state free precession imaging, which could be an alternative to the T2W TSE acquisition, with high resolution in 3D [34].

Similar to findings for other organs [19-22], we found that the overall observer variation decreased when MRI was available during delineation. Consequently, potentially smaller planning treatment volume margins can be used for CT+MRI delineations than for CT-only delineations. However, due to the exploratory nature of this research, the overall *SD* was based on a small patient group and, therefore, does not necessarily represent typical overall *SDs* for the pancreatic cancer patient population. Therefore, this study should be repeated using more patients to quantify necessary treatment margins.

In conclusion, in this exploratory study the availability of MRI images to CT during target volume delineation for pancreatic cancer improved (decreased) the overall observer variation for six out of eight structures and resulted in smaller delineated volumes compared to CT-only delineation. Yet, large local observer variations existed close to the biliary stent and/or pathological lymph nodes, indicating issues with instructions and compliance.

References

1. Yovino S, Poppe M, Jabbour S et al. Intensity-modulated radiation therapy significantly improves acute gastrointestinal toxicity in pancreatic and ampullary cancers. *Int. J. Radiat. Oncol. Biol. Phys.* 2011; 79(1):158–162.
2. Crane CH, Antolak JA, Rosen II et al. Phase I study of concomitant gemcitabine and IMRT for patients with unresectable adenocarcinoma of the pancreatic head. *Int. J. Gastrointest. Cancer* 2001; 30(3):123–132.
3. van der Horst A, Wognum S, Dávila Fajardo R et al. Interfractional position variation of pancreatic tumors quantified using intratumoral fiducial markers and daily cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 87(1):202–208.
4. van der Horst A, Lens E, Wognum S et al. Limited role for biliary stent as surrogate fiducial marker in pancreatic cancer: Stent and intratumoral fiducials compared. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):641–648.
5. Keall PJ, Starkschall G, Shukla H et al. Acquiring 4D thoracic CT scans using a multislice helical method. *Phys. Med. Biol.* 2004; 49(10):2053–2067.
6. Lens E, van der Horst A, Kroon PS et al. Differences in respiratory-induced pancreatic tumor motion between 4D treatment planning CT and daily cone beam CT, measured using intratumoral fiducials. *Acta Oncol.* 2014; 53(9):1257–1264.
7. Caravatta L, Macchia G, Mattiucci GC et al. Inter-observer variability of clinical target volume delineation in radiotherapy treatment of pancreatic cancer: A multi-institutional contouring experience. *Radiat. Oncol.* 2014; 9(1):198.
8. Yamazaki H, Nishiyama K, Tanaka E et al. Dummy run for a phase II multi-institute trial of chemoradiotherapy for unresectable pancreatic cancer: Inter-observer variance in contour delineation. *Anticancer Res.* 2007; 27(4 C):2965–2971.
9. Fokas E, Clifford C, Spezi E et al. Comparison of investigator-delineated gross tumor volumes and quality assurance in pancreatic cancer: Analysis of the pretrial benchmark case for the SCALOP trial. *Radiother. Oncol.* 2015; 117(3):432–437.
10. van Herk M. Errors and margins in radiotherapy. *Semin. Radiat. Oncol.* 2004; 14(1):52–64.
11. Rasch CRN, Steenbakkers RJHM, Fitton I et al. Decreased 3D observer variation with matched CT-MRI, for target delineation in nasopharynx cancer. *Radiat. Oncol.* 2010; 5:21.
12. Nijkamp J, de Haas-Kock DFM, Beukema JC et al. Target volume delineation variation in radiotherapy for early stage rectal cancer in the Netherlands. *Radiother. Oncol.* 2012; 102(1):14–21.
13. Versteijne E, Gurney-Champion OJ, van der Horst A et al. Considerable interobserver variation in delineation of pancreatic cancer on 3DCT and 4DCT: a multi-institutional study. Chapter 2 from this thesis.
14. Sandrasegaran K, Lin C, Akisik FM, Tann M. State-of-the-art pancreatic MRI. *Am. J. Roentgenol.* 2010; 195(1):42–53.
15. Schima W. MRI of the pancreas: tumours and tumour-simulating processes. *Cancer Imaging* 2006; 6:199–203.
16. Erturk SM, Alberich-Bayarri A, Herrmann KA et al. Use of 3.0-T MR Imaging for Evaluation of the Abdomen. *Radiographics* 2009; 29(6):1547–1563.
17. Hee SP, Jeong ML, Hei KC et al. Preoperative evaluation of pancreatic cancer: Comparison of gadolinium-enhanced dynamic MRI with MR cholangiopancreatography versus MDCT. *J. Magn. Reson. Imaging* 2009; 30(3):586–595.
18. Andreychenko A, Heerkens H, Meijer G et al. SU-E-J-07: A Functional MR Protocol for the Pancreatic Tumor Delineation. *Med. Phys.* 2014; 41(6):155–156.

19. Rasch C, Keus R, Pameijer FA et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 1997; 39(4):841–848.
20. Villeirs GM, van Vaerenbergh K, Vakaet L et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther. Onkol.* 2005; 181(7):424–430.
21. Rasch C, Barillot I, Remeijer P et al. Definition of the prostate in CT and MRI: A multi-observer study. *Int. J. Radiat. Oncol. Biol. Phys.* 1999; 43(1):57–66.
22. den Hartogh MD, Philippens MEP, van Dam IE et al. MRI and CT imaging for preoperative target volume delineation in breast-conserving therapy. *Radiat. Oncol.* 2014; 9(1):63.
23. Dalah E, Moraru I, Paulson E et al. Variability of target and normal structure delineation using multimodality imaging for radiation therapy of pancreatic cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):633–640.
24. Li J, Denniston KA, Hussain SM et al. Comparison of CT and MRI-based gross tumor volume and organ at risk delineation for pancreatic cancer patients undergoing neoadjuvant stereotactic body radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 90(1):S356.
25. Versteijne E, van Eijck CHJ, Punt CJA et al. Preoperative radiochemotherapy versus immediate surgery for resectable and borderline resectable pancreatic cancer (PREOPANC trial): study protocol for a multicentre randomized controlled trial. *Trials* 2016; 17(1):127.
26. Gurney-Champion OJ, Lens E, van der Horst A et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med. Phys.* 2015; 42(5):2638–2647.
27. Dávila Fajardo R, Lekkerkerker SJ, van der Horst A et al. EUS-guided fiducial markers placement with a 22-gauge needle for image-guided radiation therapy in pancreatic cancer. *Gastrointest. Endosc.* 2014; 79(5):851–855.
28. Barral M, Taouli B, Guui B et al. Diffusion-weighted MR imaging of the pancreas: Current status and recommendations. *Radiology* 2015; 274(1):45–63.
29. Matsuki M, Inada Y, Nakai G et al. Diffusion-weighted MR imaging of pancreatic carcinoma. *Abdom. Imaging* 2007; 32(4):481–483.
30. Steenbakkers RJHM, Duppen JC, Fitton I et al. Observer variation in target volume delineation of lung cancer related to radiation oncologist-computer interaction: a “Big Brother” evaluation. *Radiother. Oncol.* 2005; 77(2):182–190.
31. Kouwenhoven E, Giezen M, Struikmans H. Measuring the similarity of target volume delineations independent of the number of observers. *Phys. Med. Biol.* 2009; 54(9):2863–2873.
32. Hall WA, Mikell JL, Mittal P et al. Tumor size on abdominal MRI versus pathologic specimen in resected pancreatic adenocarcinoma: Implications for radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 86(1):102–107.
33. Cattaneo GM, Reni M, Rizzo G et al. Target delineation in post-operative radiotherapy of brain gliomas: Interobserver variability and impact of image registration of MR(pre-operative) images on treatment planning CT scans. *Radiother. Oncol.* 2005; 75(2):217–223.
34. Gurney-Champion OJ, Nederveen AJ, Klaassen R et al. Revisiting the potential of alternating repetition time balanced steady-state free precession imaging of the abdomen at 3 T. *Invest. Radiol.* 2016; 51(9):560–568.

Supplementary Materials A

Acquisitions

Diagnostic CT, planning CT and MRI examinations were held on separate days. On average, the diagnostic CT examination was 53 days (range 46–62 days) before the planning CT examination whereas the MRI examination was 19 days (range 15–21 days) before the planning CT examination.

The diagnostic CT scans were acquired at the referring hospitals. The scans were considered adequate for diagnostic purposes by the dedicated abdominal radiologists from our institution. The diagnostic CT scans were obtained after vascular contrast (iodine) injection and included axial scans in arterial (35 s after injection, all patients), and venous contrast phase (60 s after injection, patients 1,2 and 4) or portal contrast phase (240 s after injection, patients 1 and 4). In addition, reconstructed coronal views were available for the arterial (patients 1, 2 and 4) and venous (patients 2 and 3) contrast phases.

The planning CT scans were obtained with patients in treatment position: supine on a flat table top with hands above their head. Two planning CT scans were acquired, a 3DCT and 4DCT. The 3DCT scan was acquired during free breathing after contrast injection in the arterial contrast phase and the corresponding 3DCT images represent a snapshot of the respiratory cycle. The 3DCT was reconstructed with a slice thickness of 2.5 mm and an in-plane resolution of $1.3 \times 1.3 \text{ mm}^2$. For the 4DCT scan, the respiratory motion was monitored while data were acquired continuously during free breathing. The respiratory motion was synchronized to the CT acquisition by the respiratory gating system RPM (Real-Time Position Management, Varian Oncology Systems, Palo Alto, CA). The data were sorted according to the respiratory signal and images were reconstructed of the ten respiratory phases. Also, a maximum intensity projection (MIP) and average intensity projection (Ave-IP) were reconstructed from the images from the ten phases. The images reconstructed from the 4DCT scan had a slice thickness of 2.5 mm and an in-plane resolution of $1.0 \times 1.0 \text{ mm}^3$.

MRI scans were obtained with a 16-channel phased-array anterior coil and a 10-channel phased-array posterior coil. Four MRI scans were acquired using various sequences: T1W GE, CE T1W GE, T2W TSE and DWI (Table 3.A). The T1W GE scan obtained three echoes and used a Dixon reconstruction to obtain water only images. The T1W GE scan was acquired in a single breath-hold and images had a high isotropic resolution, allowing for a correct representation of the anatomy. However, T1W GE images show poor pancreatic tumor contrast. Therefore, an additional CE T1W GE scan was acquired after gadovist 1.0 (Bayer Healthcare, Leverkusen, Germany) administration (0.1 ml/kg; 5 ml/s, followed by a 15 ml saline flush) in the

parenchymal phase (35 seconds after injection), which improves lesion conspicuity. Due to the design of the MIPA study, another gadovist bolus had also been injected approximately 6 minutes before the CE T1W GE scan. On T2W TSE images the tumor is often poorly visible; however, its location can often be derived by finding obstructions of the pancreatic duct, which is clearly visible on these images. Also, these images can show the presence of peripancreatic inflammation. The apparent diffusion coefficient maps were generated from $b = 0$ and $600 \text{ mm}^2\text{s}$ DWI scans. DWI, and in particular, the apparent diffusion coefficient maps from DWI, give information about the diffusivity of tissue. In general, pancreas carcinomas show increased cellular density, causing decreased diffusivity [28, 29].

Table 3.A. MRI sequence settings.

	T1W GE	T2W TSE	DWI
Resolution (mm ²)	1.7 × 1.7	1.3 × 1.6	3 × 3
Slice thickness (mm)	1.7	5	3.7
FOV RL × AP × FH (mm ³)	400 × 353 × 95	400 × 369 × 145	432 × 108 × 67
TR/TE/ΔTE (ms)	4.7/1.26/0.9	779/80/–	> 2200/45/–
Parallel imaging	2 (AP)/1.5 (FH)	2 (AP)	1.3 (AP)
Fat saturation	Dixon	SPAIR	SPIR and gradient reversal
Flip angle (°)	9 (25 post contrast)	90 (120 refocussing)	90
Respiratory compensation	1 breath-hold	3 breath-holds	Respiratory triggering using a navigator

Abbreviations: T2W TSE = T2-weighted turbo spin echo; DWI = diffusion-weighted MRI; FOV = field of view; RL = right–left; AP = anterior–posterior; FH = foot–head; TR = repetition time; TE = echo time; ΔTE = increase in TE; SPAIR = spectral attenuated inversion recovery; SPIR = spectral presaturation with inversion recovery.

Supplementary Materials B

Lymph nodes, stents/drain, and markers

Table 3.B. Number of observers who delineated the lymph nodes.

lymph nodes	3DCT+MRI	3DCT-only*	4DCT+MRI	4DCT-only*
patient 2, portocaval	2	5	4	4
patient 2, along common hepatic duct	2	3	3	2
patient 4	2	2	1	2

* These results are from the CT-only study [13].

Table 3.C. Number of observers who delineated the stent or drain.

Patients	3DCT+MRI	3DCT-only*	4DCT+MRI	4DCT-only*
1	2	3	2	5
2	1	2	2	1
3	4	6	5	4
4	8	8	8	7

Number of observers who delineated at least 50% of the stent/drain in at least three slices are reported

* These results are from the CT-only study [13].

Table 3.D. Markers located within the delineated volumes.

	3DCT+MRI		3DCT-only*		4DCT+MRI		4DCT-only*	
1	24/24	(100%)	24/24	(100%)	24/24	(100%)	23/24	(100%)
2 [†]	1/24	(4%)	5/24	(21%)	5/24	(21%)	4/24	(17%)
3	11/24	(46%)	22/24	(92%)	14/24	(58%)	16/24	(67%)
4	24/24	(100%)	24/24	(100%)	24/24	(100%)	21/24	(88%)

The number of volumes including the marker over the number of markers multiplied by number of observers (3 × 8 = 24). Markers that were included partially in the delineated volume were counted as delineated.

* These results are from the CT-only study [13].

[†] In patient 2, the markers were mistakenly implanted outside the tumor.

Supplementary Materials C

Local standard deviation

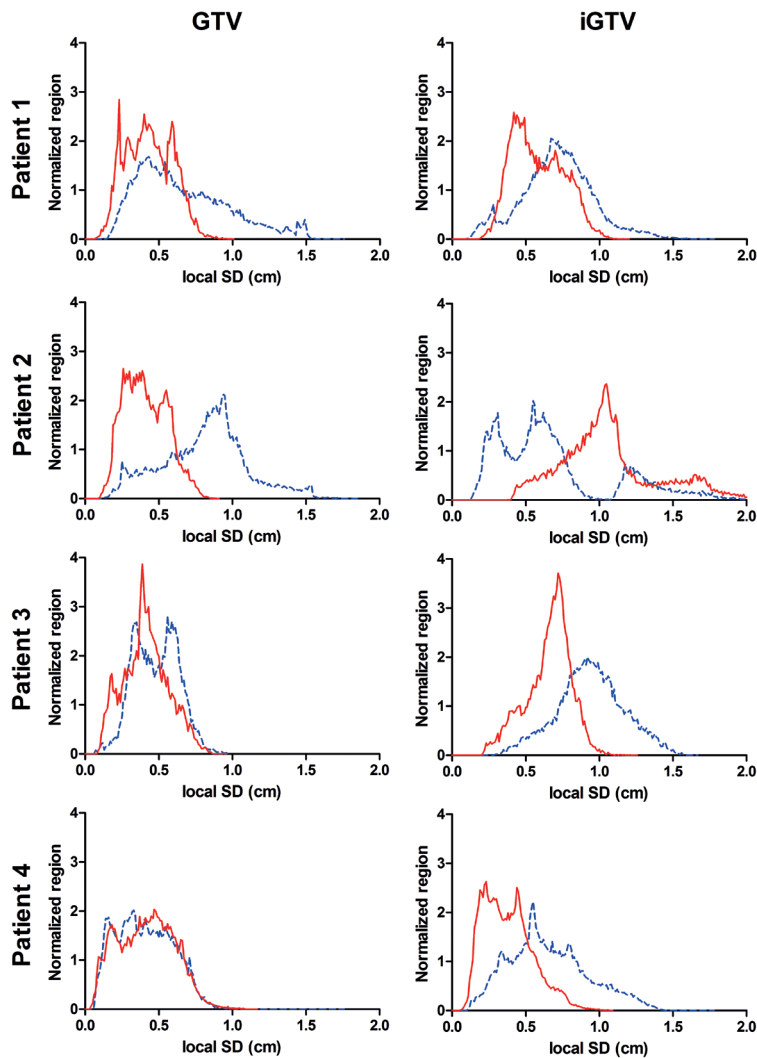
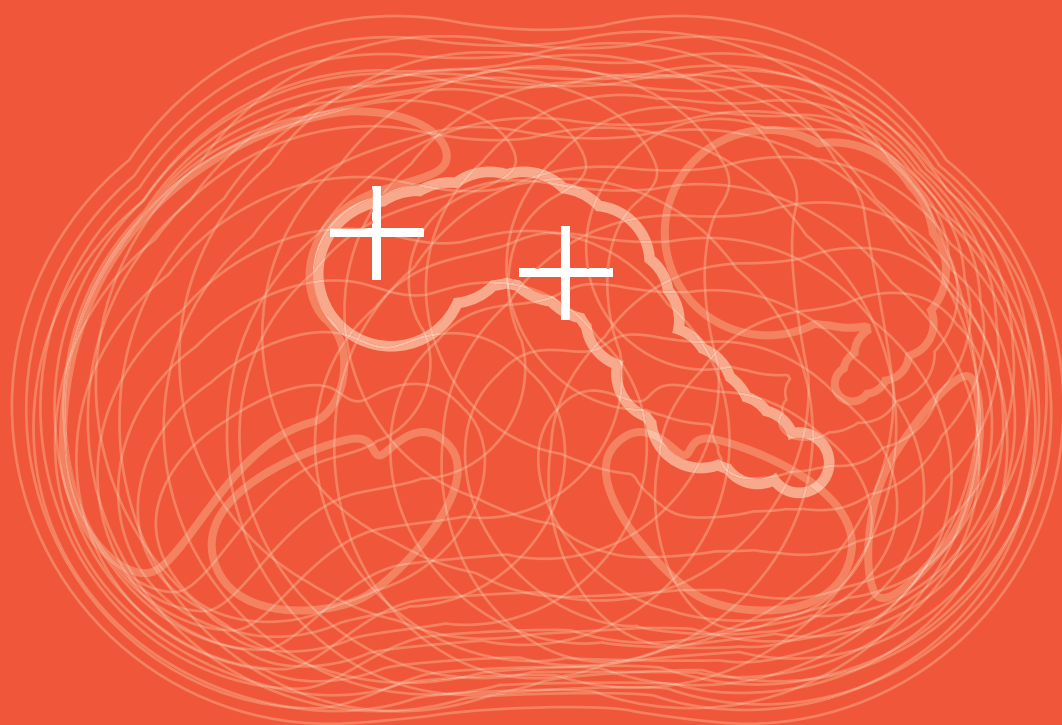


Figure 3.A. Normalized histogram of the local SD of the CT-only study (dotted gray line) and CT+MRI study (solid black line). Graphs are based on the local SDs from on average 32,324 (range 9,728–66,504) points on the median surface and normalized to the total surface of the median surface.



CHAPTER 4

Visibility and artifacts of fiducial markers

Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI

Oliver J Gurney-Champion

Eelco Lens

Astrid van der Horst

Antonetta C Houweling

Remy Klaassen

Jeanin E van Hooft

Jaap Stoker

Geertjan van Tienhoven

Aart J Nederveen

Arjan Bel

Medical Physics 2015;42(5):2638–2647

Abstract

Purpose: In radiation therapy of pancreatic cancer, tumor alignment prior to each treatment fraction is improved when intratumoral gold fiducial markers (from here onwards: markers), which are visible on computed tomography (CT) and cone beam CT, are used. Visibility of these markers on magnetic resonance imaging (MRI) might improve image registration between CT and magnetic resonance (MR) images for tumor delineation purposes. However, concomitant image artifacts induced by markers are undesirable. The extent of visibility and artifact size depend on MRI-sequence parameters. The authors' goal was to determine for various markers their potential to be visible and to generate artifacts, using measures that are independent of the MRI-sequence parameters.

Methods: The authors selected ten different markers suitable for endoscopic placement in the pancreas and placed them into a phantom. The markers varied in diameter (0.28–0.6 mm), shape, and iron content (0%–0.5%). For each marker, the authors calculated $T2^*$ -maps and $\Delta B0$ -maps using MRI measurements. A decrease in relaxation time $T2^*$ can cause signal voids, associated with visibility, while a change in the magnetic field $B0$ can cause signal shifts, which are associated with artifacts. These shifts inhibit accurate tumor delineation. As a measure for potential visibility, the authors used the volume of low $T2^*$, i.e., the volume for which $T2^*$ differed from the background by > 15 ms. As a measure for potential artifacts, the authors used the volume for which $|\Delta B0| > 9.4 \times 10^{-8}$ T (4 Hz). To test whether there is a correlation between visibility and artifact size, the authors calculated the Spearman's correlation coefficient (R_s) between the volume of low $T2^*$ and the volume of high $|\Delta B0|$. The authors compared the maps with images obtained using a clinical MR-sequence. Finally, for the best visible marker as well as the marker that showed the smallest artifact, the authors compared the phantom data with *in vivo* MR-images in four pancreatic cancer patients.

Results: The authors found a strong correlation ($R_s = 1.00$, $p < 0.01$) between the volume of low $T2^*$ and the volume with high $|\Delta B0|$. Visibility in clinical MR-images increased with lower $T2^*$. Signal shift artifacts became worse for markers with high $|\Delta B0|$. The marker that was best visible in the phantom, a folded marker with 0.5% iron content, was also visible *in vivo*, but showed artifacts on diffusion-weighted images. The marker with the smallest artifact in the phantom, a small, stretched, ironless marker, was indiscernible on *in vivo* MR-images.

Conclusion: Changes in $T2^*$ and $\Delta B0$ are sequence-independent measures for potential visibility and artifact size, respectively. Improved visibility of markers correlates strongly to signal shift artifacts; therefore, marker choice will depend on the clinical purpose. When visibility of the markers is most important, markers that

contain iron are optimal, preferably in a folded configuration. For artifact sensitive imaging, small ironless markers are best, preferably in a stretched configuration.

Introduction

Radiation therapy aims at delivering a high radiation dose to the tumor while minimizing the dose to healthy tissues. For pancreatic cancer, this is challenging as the radiation dose to the tumor is limited due to the surrounding vital organs. Moreover, pancreatic tumors show both intra-fractional motion and interfractional position variations [1, 2]. Furthermore, these tumors are often poorly defined on computed tomography (CT) images.

To correct for daily position variations of the tumor, the target position can be verified by comparing in-room cone beam computed tomography (CBCT) images with the planning CT scan. However, soft-tissue contrast is low on CBCT images and, consequently, CBCT to CT registrations are often based on bony anatomy, hence, remaining inter-fractional position variations of the tumor can still persist.

Radiation may be delivered more accurate when position verification is done based on the tumor position rather than the bony anatomy. Implanted intratumoral fiducial gold markers (from here onwards: markers) that are visible on CT and CBCT, enable this [1]. To ensure endoscopic placement of markers in the pancreas through a 19-gauge needle, the markers require a diameter (\varnothing) of less than 0.686 mm [3]. Though, markers that fit in a 22-gauge needle (markers with $\varnothing < 0.413$ mm) may be preferred, as this needle is more maneuverable [4].

Due to poor soft-tissue contrast, tumor delineation on CT-images can result in large inter-observer variations [5]. Magnetic resonance imaging (MRI), shown to improve tumor delineation in various tumor sites [6-8], may also benefit pancreatic cancer patients. Also, MRI can be used for detecting and characterizing pancreatic tumors using contrast enhanced MRI or diffusion-weighted MRI (DWI) [9, 10].

When both MRI and markers are used, visibility of the markers on magnetic resonance (MR) images might improve CT-MR or CBCT-MR image registration [11, 12]. Several larger markers ($\varnothing \geq 1$ mm) are proven to be visible on MRI [11-13]. However, smaller markers, such as the ones implanted in pancreatic cancer patients, may lack visibility and, therefore, prohibit marker-based registration. For example, Chan et al. argue that from their selection of markers ($0.5 \leq \varnothing \leq 1.2$ mm), the larger markers should be chosen when visibility on MRI is desired [14]. In addition, markers may cause image artifacts, in particular in DWI [15]. On diffusion-weighted (DW) images of the prostate, some markers ($\varnothing = 1$ mm, length = 3 mm) produced artifacts up to

25 mm in diameter which also resulted in an inaccurate apparent position [16]. This prevents accurate tumor delineation close to the markers. Therefore, selecting the optimal marker is important and the knowledge about marker visibility and artifacts the marker causes is crucial in this selection.

Both the visibility and extend of the artifact depend on the choice of MRI-sequence parameters and therefore, findings are hard to generalize. In this paper, we propose an innovative sequence-independent framework for describing the effects of markers on MR-images by quantifying their effects on the magnetic field, B_0 , and on the relaxation time, T_2^* [17-19]. Changes in B_0 , occurring due to differences in magnetic susceptibility of the marker compared to the surrounding tissue, cause signal shifts. Signal shifts result in bright and dark spots in the image and changes in the apparent position of the marker. These signal shifts are therefore associated with artifacts. For a given change in B_0 , the amount of shift depends only on the bandwidth (BW) of the MRI-sequence and thus, artifacts can be reduced by increasing the sequence BW . In addition, markers decrease the local T_2^* , thereby creating the potential for a signal void. These signal voids are associated with visibility. For a given decrease in T_2^* , the amount of signal loss in the vicinity of the marker depends strongly on the echo time (TE) of the sequence; therefore, the visibility can be tuned by adjusting the TE .

The goal of this study was to determine the visibility of several markers as well as the artifacts they generate on MR-images. We investigated whether there is a relation between visibility (signal voids) and artifact size (signal shifts). This was done by quantifying their propensity to generate signal voids and signal shifts in a phantom. The results were compared to the measurements in pancreatic cancer patients.

Materials and Methods

Fiducial gold markers

We selected ten markers, of which six Gold Anchors (Naslund Medical AB, Huddinge, Sweden), two Visicoils (IBA Dosimetry, Schwarzenbruck, Germany), one PointCoil, and one FlexiCoil (both CIVCO Medical Solutions, Kalona IA). All markers had a diameter $\varnothing < 0.686$ mm, such that they fit in 19-gauge or 22-gauge needles. All markers consisted of at least 99.5% gold. Markers 1 and 2 contained some iron to improve their visibility on MR-images [20]. The markers varied in size, shape, and composition (Table 4.1, Fig. 4.1). All Gold Anchors were analyzed in a stretched and a folded configuration.

To mimic our clinical practice, we manually cut down the markers that were longer than 12 mm (markers 1, 2, and 5) to a length of approximately 10 mm. Then, for all markers, we measured their length with a ruler (estimated absolute error of ≤ 1 mm).

Table 4.1. Properties of ten investigated markers.

	Diameter (mm)	Measured length (mm)	Composition (%)	Product name	Shape
1	0.28	10	99.5 \pm 0.05 Au, 0.5 \pm 0.05 Fe	Gold Anchor	NL
2	0.28	10	99.5 \pm 0.05 Au, 0.5 \pm 0.05 Fe	Gold Anchor	Folded NL
3	0.28	10	> 99.99 Au	Gold Anchor	NL
4	0.28	10	> 99.99 Au	Gold Anchor	Folded NL
5	0.35	12	> 99.95 Au	Visicoil	CL
6	0.40	9	> 99.99 Au	Gold Anchor	NL
7	0.40	12	> 99.99 Au	Gold Anchor	Folded NL
8	0.50	10	> 99.95 Au	Visicoil	CL
9	0.6	5	> 99.99 Au	PointCoil	Helical coil
10	0.6	10	> 99.99 Au	FlexiCoil	CL with solid endpoints

Note: Compositions not mentioned were unspecified by the manufacturer.

Abbreviations: NL = notched linear; CL = coiled linear (see Fig. 4.1 for examples)

Phantom study

We filled ten jars (radius = 25 mm, height = 70 mm) halfway with agarose gel. The gel contained 1.3 g NaCl, 85 μ mol GdCl₃, 25 g agarose, and 1.5 g sodium benzoate, per liter demi-water, mimicking pancreatic tissue relaxation properties [21, 22]. Once the agarose gel had set, we placed each marker at the center of a jar on top of the gel and filled up the rest of the jar with the remaining gel.

The jars were placed in a jar rack that fixed their positions, which was placed in a container filled with water to minimize air interfaces close to the markers (Fig. 4.2). Such interfaces could induce additional, unwanted, changes in ΔB_0 .

Imaging

We obtained CT images with a LightSpeed RT 16 CT-scanner (General Electric Company, Waukesha WI), with 1.0×1.0 mm² in-plane resolution and a 2.5 mm slice thickness. Subsequently, CBCT images were acquired using the on-board kV-imager of the Synergy linear accelerator (Elekta Oncology systems, Crawley, UK) with

$1.0 \times 1.0 \times 1.0 \text{ mm}^3$ voxels. The CT and CBCT images are meant for illustration and no special attention was paid to the alignment.

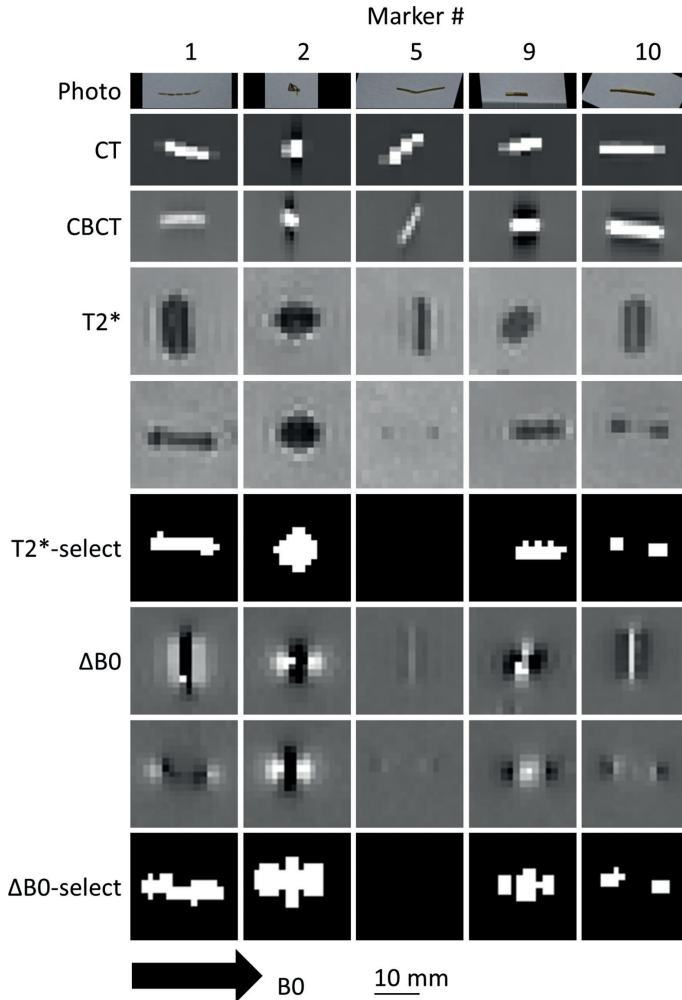


Figure 4.1. Example images for each of the marker shapes. From left to right: notched linear (marker 1), folded notched linear (marker 2), coiled linear (marker 5), helical coil (marker 9), and coiled linear with solid endpoints (marker 10). From top to bottom, the figure shows different imaging modalities. Photos of the markers were obtained using a digital camera. MR-images were acquired with the markers perpendicular (top) and aligned (bottom) with respect to the B₀-field (arrow). Photo, CT, CBCT, and MR-images were taken at different times and therefore marker rotation between modalities occurred. The markers seem longer on CT due to the reconstruction method. Marker 5 was slightly bent when cut down (photo) but bent back to its original stretched configuration in the phantom. The T2*-selected and ΔB0-selected images use the aligned case as an example to illustrate in white which voxels contributed to the volume of low T2* and ΔB0, respectively. For marker 5, no voxels were selected in this scan.

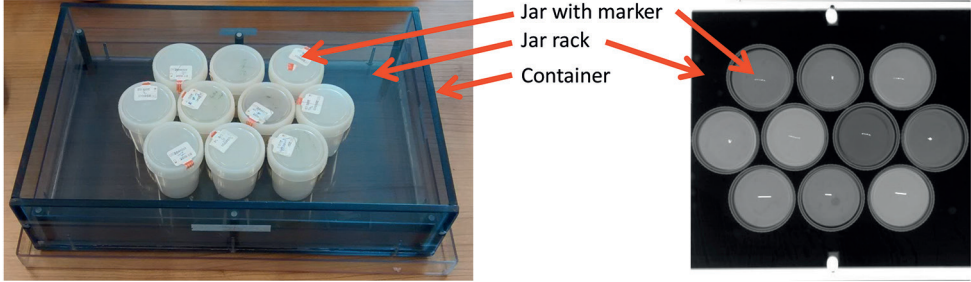


Figure 4.2. Photo and x-ray radiograph of the phantom. On the x-ray, the markers are visible as white stripes and dots in the jars, depending on the configuration of the marker.

We acquired all MR-images on an Ingenia 3 T scanner (Philips Healthcare, Best, the Netherlands), using a 16-channel phased-array anterior coil and a 10-channel phased-array posterior coil. For each sequence, we obtained four MR-images: two different marker orientations (aligned and perpendicular to B_0 for the linear markers) as well as two different read-out directions [left–right (LR) and feet–head (FH)]. As we will investigate whether there is a difference between the two marker orientations on MRI, marker alignment for the MRI was more critical, and was done within $\pm 15^\circ$. We expect little effect from rotations within this limit.

To produce maps showing the local T_2^* and ΔB_0 values per voxel, a T_2^* -map and ΔB_0 -map, respectively, we obtained a T1-weighted eight echo steady state free precession (SSFP) MR-images of the phantom (Table 4.2). These images were reconstructed at a 1.4 mm isotropic resolution.

We created T_2^* -maps from the magnitude data of the SSFP image by fitting

$$S(TE) = S(0) e^{\frac{TE}{T_2^*}} \quad (4.1)$$

to the signal, $S(TE)$, as function of echo time, TE , for each voxel. Note that a decrease in signal due to the lack of hydrogen atoms in the marker, and a decrease in signal due to RF shimming, are already present at $TE = 0$ ms and thus are taken into account by $S(0)$; therefore, they do not effect T_2^* . We used a phase unwrapping algorithm to calculate ΔB_0 from the phase data from the same SSFP image [23, 24]. To remove marker-unrelated low frequency fluctuations in these ΔB_0 -maps, we subtracted a low-pass filtered version of the ΔB_0 -map (3D Gaussian, standard deviation: four voxels in each direction) from the ΔB_0 -maps.

To determine whether the voids and signal shifts on clinical sequences were in agreement with the findings from the T_2^* -maps and ΔB_0 -maps, respectively, we obtained MR-images of the phantom using the imaging sequences from our clinical

protocol that have the highest potential to show the marker or to show artifacts (Table 4.2). The Dixon images were acquired using a multi-echo T1-weighted FFE sequence and used a Dixon reconstruction to generate a water-only image. These images are particularly suitable to visualize anatomical features. For the multi-echo SSFP sequence, the visibility of a marker improves at later echoes; therefore, it allows us to tune the visibility. The DWI sequence is the most sensitive to artifacts.

The DWI sequence in the protocol ($DWI_{\text{high}} BW$) was configured such that it had a high BW/voxel to minimize marker signal shift artifacts. For the phantom study, we also acquired a DWI sequence ($DWI_{\text{low}} BW$) using a lower BW/voxel for comparison. Both DWI sequences had a diffusion weighting of 0 and 600 mm^2/s , rest slabs to prevent fold-over artifacts, a spectral presaturation with inversion recovery (SPIR) pulse for fat saturation, and gradient reversal for off-resonance suppression during slice excitation.

Patient study

We obtained CT, CBCT (settings similar as for the phantom study), and MR-images (clinical protocol, Table 4.2) of four patients with markers (two males, two females, aged between 62 and 80, mean age 69). The patients were diagnosed with borderline resectable or resectable pancreatic head carcinomas. They received radiochemotherapy prior to explorative laparotomy in the framework of the multi-center PREOPANC trial (EudraCT No. 2012-003181-40) and they received a MRI as part of the translational MIPA study (Clinicaltrials.gov no. NCT01989000). The ethics committee of the Erasmus Medical Center and the Academic Medical Center, respectively, approved the protocols and all patients gave informed consent to both. The patients received intratumoral markers as part of our standard clinical practice. Feasibility of the marker implantation procedure was shown in an earlier study [4].

Patient 1 received two $\varnothing = 0.28$ mm Gold Anchors with 0.5% iron, which had the largest signal void in the phantom experiment, and one $\varnothing = 0.35$ mm Visicoil, which had the smallest signal shift, and patients 2–4 received three $\varnothing = 0.35$ mm Visicoils. Patients 1–3 had metallic biliary stents.

Data analysis

As stated in the Introduction, marker visibility is generated by signal voids induced by the marker. The potential for a marker to generate these voids depends on the decrease in $T2^*$ it causes. Given a $T2^*$ -map, the size of the voids, which strongly depends on the TE of the sequence, can be determined for each MRI-sequence. It is,

however, impossible to quantify and report this for every MR-sequence. Therefore, we ranked the visibility in two ways: the volume of low $T2^*$ (which is related to the potential size of the signal voids) and the decrease in $T2^*$ relative to the background (which is related to the potential magnitude of the signal voids).

Table 4.2. Sequence parameters used for the phantom and *in vivo* experiments.

	Phantom study		Clinical protocol (phantom and patients)		
	SSFP	DWI _{low BW}	Dixon	SSFP	DWI _{high BW}
FOV (mm ³) (LR × AP × FH)	400 × 350 × 90	350 × 100 × 90	400 × 350 × 95	400 × 350 × 95	430 × 110 × 70
Resolution (mm ²)	1.4 × 1.4	2.0 × 2.0	1.7 × 1.7	2.3 × 2.3	3.0 × 3.0
Slice thickness (mm) (acquired/ reconstructed)	2.8/1.4	2.0	1.7	4.6/2.3	3.7
TR/TE/ΔTE (ms)	32.9/2.3/2.3	3500/50/—	4.7/1.15/1.0	20/2.3/2.3	4000/45/—
Parallel imaging	—	1.7	2/1.5 (AP/FH)	2/1.5 (AP/FH)	1.7
Half Fourier	—	0.8	—	—	0.8
Flip angle (°)	12	90	8	12	90
Averages	2	8	1	1	15
<i>In vivo</i> respiratory compensation	—	—	Breath hold	Breath hold	Navigator triggered
BW/voxel (Hz)	1252	17.5 ^a	1602	1973	58.8 ^a
Scan time (s)	470	126	20	23	148 ^b

^a In the phase direction.

^b The total scan time without respiratory triggering.

Abbreviations: TR = repetition time; ΔTE = increase in TE; AP = anterior–posterior direction; FA = flip angle.

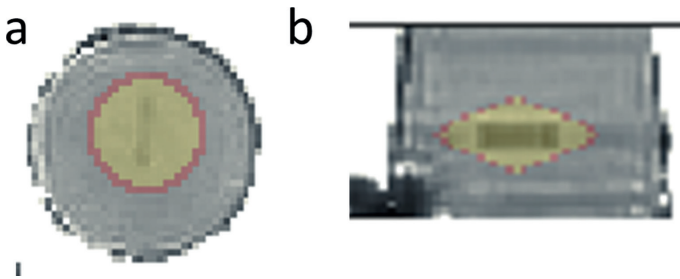


Figure 4.3. Example of how the ellipsoidal ROIs were placed on the $T2^*$ -maps. Figures (a) and (b) show coronal and sagittal view of the ellipsoid, respectively. The inner ROI indicates the region in which voxels were counted and the shell indicates the 1-voxel wide shell used to determine background values of $T2^*$.

For each marker, we selected an ellipsoidal region of interest (ROI) (Fig. 4.3) (485–1713 voxels). We visually checked that the ROI included the entire extent of the $T2^*$ as well as $\Delta B0$ effects and avoided air bubbles and effects from the water–gel/water–air boundaries at the edge, top, and bottom of the jar.

As a measure of the volume of low $T2^*$, we used the volume of all voxels in our ROI (Fig. 4.3) with a $\Delta T2^* = T2^*_{\text{background}} - T2^*_{\text{min}}$ larger than 15 ms. Here, $T2^*_{\text{background}}$ is the average $T2^*$ of all voxels in a 1-voxel wide shell around the ROI (287–648 voxels) (Fig. 4.3). We visually checked that the ROI was not contaminated by effects of the marker, nor air bubbles and effects from the water–gel/water–air boundaries at the edge, top and bottom of the jar. The cut-off value of 15 ms was chosen in order to differentiate markers according to the marker's volume of low $T2^*$. As a measure of the decrease in $T2^*$ relative to the background, we calculated:

$$frac_{\Delta T2^*, \max} = \frac{T2^*_{\text{background}} - T2^*_{\text{min}}}{T2^*_{\text{background}}} \quad (4.2)$$

Here, $T2^*_{\text{min}}$ is the mean $T2^*$ of the three voxels with lowest $T2^*$ within the ROI.

In addition to voids, markers also generate artifacts by inducing signal shifts. The potential for a marker to generate such shifts depends on the changes in $\Delta B0$ it creates. Given a $\Delta B0$ -map and the sequence BW , the distance signal shifts can be determined for each MRI-sequence. Similar to $T2^*$, we rank artifacts in two ways: the volume of affected $\Delta B0$ (which is related to the potential size of the artifact) and the maximum difference in $\Delta B0$ (which is related to the potential magnitude of the artifact).

As a measure for the volume of affected $\Delta B0$, we used the volume of all voxels within the ROI with a value of $|\Delta B0| > 9.4 \times 10^{-8}$ T (4 Hz). The value of 9.4×10^{-8} T (4 Hz) allowed us to differentiate between markers according to the marker's volume of high $|\Delta B0|$. The absolute value is taken as the marker can either increase or decrease $B0$. In addition, as a measure for maximum difference in $\Delta B0$, $\Delta_{\max} B0$, we determined the difference between the mean $\Delta B0$ of the three voxels with highest $\Delta B0$ and the mean $\Delta B0$ of the three voxels with lowest $\Delta B0$ within the ROI.

We compared the volume with low $T2^*$ and with large $|\Delta B0|$, $frac_{\Delta T2^*, \max}$, and $\Delta_{\max} B0$ between both marker orientations. The orientation of a marker in vivo with respect to read-out direction as well as to $B0$ is random and cannot be fixed at will. Therefore, for each marker, we averaged the volumes of low $T2^*$, the volumes of high $|\Delta B0|$, the $frac_{\Delta T2^*, \max}$ and the $\Delta_{\max} B0$ over both orientations of the markers with respect to $B0$ and for both read-out directions, for all data analyses. To test whether there was a

correlation between the visibility and the artifact size, a Spearman rank correlation coefficient, R_s , was calculated between the volume with low $T2^*$ and the volume with large $|\Delta B0|$, using SPSS (version 21, IBM, New York). R_s was also calculated between $\Delta_{\max} B0$ and $frac_{\Delta T2^*, \max}$.

To illustrate how our data can be used, we calculated for each marker the expected maximum signal shift, using a BW per voxel (BW_{voxel}) of 17.5 ms (as for $DWI_{\text{low } BW}$) and the measured $\Delta_{\max} B0$,

$$\Delta x_{\#, \max} = \frac{\gamma \times \Delta_{\max} B0}{2\pi \times BW_{\text{voxel}}} \quad (4.3)$$

Here, γ is the gyromagnetic ratio. Also, this equation was used to estimate the maximum voxel shift in our $T2^*$ - and $B0$ -maps by taking the largest measured $\Delta_{\max} B0$ and the BW from the $T1$ -weighted SSFP measurement (Table 4.2).

To determine visibility of the markers *in vivo*, an experienced abdominal radiologist (20 yr experience) examined the MR-images.

Results

Phantom study

All markers were visible on both CT and CBCT. The markers appeared larger on CT than in the photo and on CBCT due to partial volume effect.

Figure 4.1 shows for five markers the $T2^*$ -maps and $\Delta B0$ -maps. One of the markers showed for the FH read-out direction a low $T2^*_{\text{background}}$ (39.2 ms) compared to the other markers; therefore, those two data points (both marker orientations), were excluded from the analyses. The $T2^*_{\text{background}}$ values varied between 43.3 and 47.0 ms (mean: 45.1 ms). In the same figure we showed binary maps illustrating in white which voxels contributed to the volume of low $T2^*$ or to the volume with high $|\Delta B0|$ ($T2^*$ -selected and $\Delta B0$ -selected, respectively).

The seven stretched markers had a larger volume of low $T2^*$ ($2.6 \times$ larger on average), a larger volume of high $|\Delta B0|$ ($2.5 \times$), a higher $frac_{\Delta T2^*, \max}$ ($1.4 \times$) and a larger $\Delta_{\max} B0$ ($2.3 \times$) when oriented perpendicular to $B0$ compared to the parallel orientation.

There was a strong non-linear correlation ($R_s = 1.00$, $p < 0.01$) between volume with low $T2^*$ and volume with high $|\Delta B0|$ (Fig. 4.4 a). Also $frac_{\Delta T2^*, \max}$ and $\Delta_{\max} B0$ (Fig. 4.4 b) were strongly correlated ($R_s = 0.94$, $p < 0.01$). The non-linearity of the correlation was

due to the fact that $0 \leq \text{frac}_{\Delta T^*, \max} \leq 1$, by definition, whereas $\Delta_{\max} B0$ can, in theory, increase unlimited. All Gold Anchor pairs with different configurations (markers 1–2, 3–4, and 6–7) had a larger volume with low $T2^*$, volume with high $|\Delta B0|$, $\text{frac}_{\Delta T^*, \max}$ and $\Delta_{\max} B0$ for the folded configuration compared to the unfolded configuration (Fig. 4.4).

Larger changes in $T2^*$ -maps correlated with larger signal voids on the Dixon images and the multi-echo SSFP images (Fig. 4.5). However, the apparent void size differed between the sequences, depending on the $T2^*$ weighting of the derived images.

Larger changes on the $\Delta B0$ -maps correlated with larger signal shift artifacts in DW-images (Fig. 4.5). The calculated maximum distances voxels could theoretically shift with respect to each other, given the measured $\Delta_{\max} B0$ and a BW of 17.5 Hz were 9.8 mm for marker 2, ≤ 0.7 mm for markers 3 and 5, and between 1.7 and 6 mm for the rest (Table 4.A from supplementary materials). For $DWI_{\text{high } BW}$, the BW was 3.4 times larger and voxels were 1.5 times larger than for $DWI_{\text{low } BW}$; therefore, the signal is expected to shift 0.45 times as far. The largest measured value of $\Delta_{\max} B0$ corresponded to a maximum signal shift in our $T2^*$ - and $B0$ -maps of 0.07 voxels (Eq. 4.3); this effect was ignored in the data analysis.

Patient study

The radiologist found that marker 2 was visible *in vivo* on most images and of comparable size as in the phantom images (Figs. 4.6 a and c). In the Dixon image, one out of the two markers 2 was indiscernible. However, it was visible on the SSFP-image. Both markers were indiscernible from anatomical features in the DW-images.

The radiologist could not discern marker 5, the marker with the smallest amount of signal shift in our phantom, on either of the MR-images of any of the four patients (Fig. 4.6). Patients 1–3 had a biliary stent, recognizable as a hyperdense circle with a hypodense center on the CT images. In the Dixon image (Fig. 4.6 b), a spot is visible that might be interpreted as a marker, but in fact was a duct.

Discussion

Our study is the first to demonstrate marker visibility as well as artifact properties on MRI using sequence-independent measures: $\Delta B0$ and $T2^*$. Also, our study is the first to investigate the behavior on MRI of a range of markers that are small enough to fit in the 19 or 22-gauge needles that are used for endoscopic placement in the pancreas. We found that markers that yielded large volumes of low $T2^*$ also produced large

volumes of high $|\Delta B_0|$, indicating that the potential marker visibility and potential for the marker to generate artifacts are related. The marker that was best visible in the phantom was detectable *in vivo*, whereas the marker that yielded the smallest signal shifts was indiscernible *in vivo*. For the remaining markers, the *in vivo* visibility and signal shifting can be estimated from their T_2^* and ΔB_0 properties (Fig. 4.4), either using Eqs. (1) and (3), or by comparing them to the two markers that were imaged *in vivo*.

The method we have presented in this study can be used as a sequence-independent measure to compare all fiducial markers that have a T_2^* based visibility.

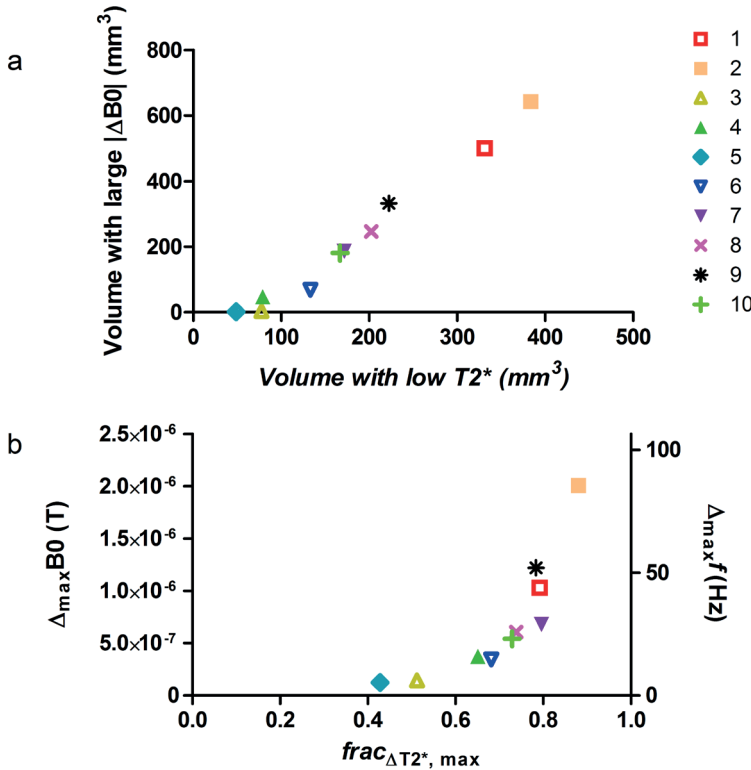


Figure 4.4. (a): The volume of affected ΔB_0 (a measure for the potential size of the artifact) plotted against the volume of low T_2^* (a measure for the potential size of the signal voids). (b): the maximum difference in ΔB_0 (a measure for the potential magnitude of the artifact) plotted against T_2^* relative to the background (a measure for the potential magnitude of the signal voids). On the right axis, the $\Delta_{\max} B_0$ is given as the frequency difference of the proton spins, $\Delta_{\max} f$. Both graphs give values averaged over the two orientations and readout directions; a graph with all separate measurements is shown in the supplementary materials.

These markers can be characterized by the two measures we introduced for potential visibility (volumes of low $T2^*$ and $\frac{1}{\Delta T2^*, \max}$) and the two measures we introduced for potential level of the artifact (volume of high $|\Delta B0|$ and $\Delta_{\max} B0$). The $T2^*$ -maps and $\Delta B0$ -maps can also be useful as a sequence-independent tool to assess the potential visibility and the potential to generate image distortions for other forms of implants.

Besides $B0$ and $T2^*$, there are additional effects that can cause artifacts or marker visibility. First of all, the markers contain no hydrogen atoms and therefore should generate local signal voids. Considering the largest marker in our study, which had a diameter of 0.6 mm and a length of 10 mm, and the high resolution at which we scanned our phantom, $1.4 \times 1.4 \times 2.8 \text{ mm}^3$, the volume of the entire marker is approximately half the volume of a single voxel. In reality, the marker will be spread out over multiple voxels and partial volume effects will occur. Therefore, this effect is negligible. Second, switching gradients could induce eddy currents in the markers and cause local signal voids [25]. Finally, markers may induce RF pulse shielding, which can produce artifacts [26-28]. Due to the small size of the markers, both the gradient induced eddy current artifacts and the RF shielding artifacts are expected to be secondary to the $T2^*$ decay effects and $B0$ effects for clinical MR-sequences [18, 25]. We have not seen any evidence of such artifacts during our study and, therefore, we believe that measuring the static $B0$ and $T2^*$ is sufficient to describe the primary effects of markers on clinical MR-images. Nevertheless, we would like to point out that for some image-sequences and/or markers additional artifacts may arise through these mechanisms.

Visibility of markers is based on a negative contrast. Hence, in areas where the background signal was low, it was harder to detect the markers. This is illustrated in the Dixon image of patient 1 (Fig. 4.6 a: Dixon). The marker that was not visible on the Dixon image was visible in the SSFP image, due to the different contrasts of the image. We therefore believe that it is important that multiple MR-sequences are obtained, yielding separate contrasts, when markers detection is desired.

Marker orientation contributes substantially to the markers' potential visibility and its potential to cause artifact: a folded marker configuration is preferable when visibility has priority and a stretched configuration is preferable when undisturbed DW-imaging is desired. However, in a clinical situation, it can be hard to ensure a specific marker orientation when implanting. In our experience, Gold Anchor markers tend to fold during implantation.

One of the limitations of our study is that we only focused on 3 T MRI [9, 10]. Currently, MRI of the pancreas is done mainly on 1.5 and 3 T systems. However, 3 T offers a superior signal-to-noise ratio when compared to 1.5 T and will most probably become the system of choice for pancreatic MRI in the future [29, 30]. Therefore,

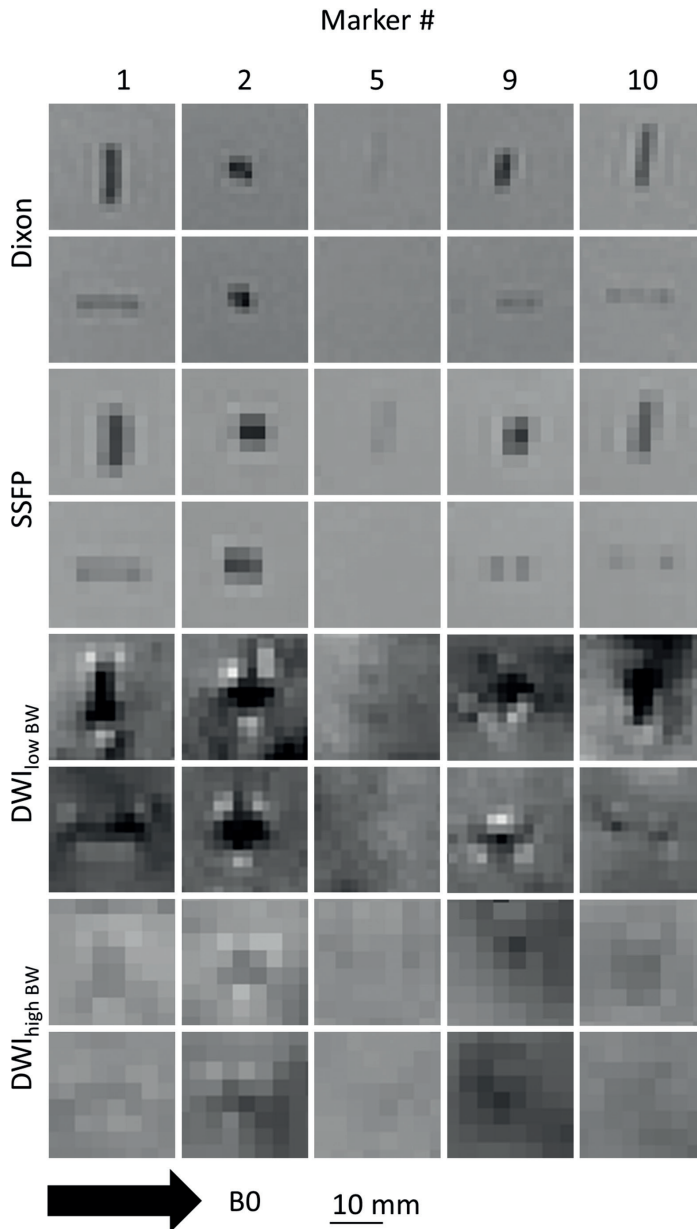


Figure 4.5. Images of the same markers as in Fig. 4.1 (left to right), obtained using different MRI protocols (top to bottom). The B0-field direction was in the plane of the images (arrow). The MR-images were acquired with markers perpendicular (**left**) and aligned (**right**) with respect to the B0-field (arrow). The SSFP image shows the 3rd echo. The DW-images had a diffusion weighting of 0 mm²s.

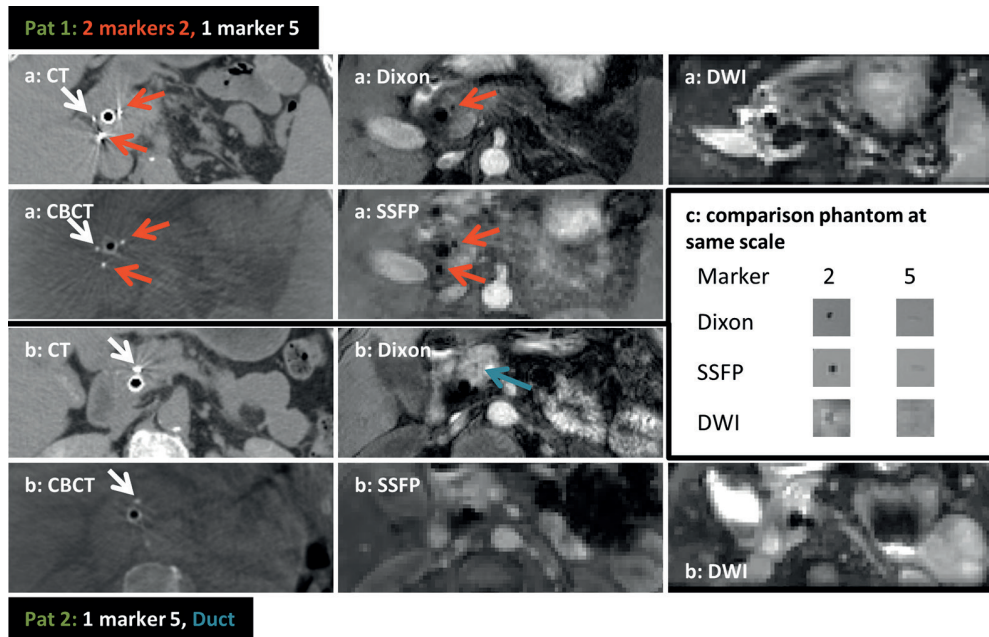


Figure 4.6. CT, CBCT, and MR-images of markers in two patients. (a): Patient 1 with two markers 2 and one marker 5. (b): Patient 2 had three markers 5 of which only one is visible on these CT and CBCT slices. The spot is a duct, not a marker. (c): MR-images of the same markers in the phantom using the same magnification as for the *in vivo* images.

this study focussed on 3 T MRI. Generally, signal shifts and signal voids increase with magnetic field strength. It was shown for comparable gold fiducial markers that visibility and artifact size indeed increased with higher field strength [14, 16].

In addition, our study included only four patients; however, in our phantom study, we were able to characterize many markers in a controlled environment. The limited patient data enabled us to show that markers had the same properties *in vivo*.

Another limitation is that the chosen cut-off values for volume of low $T2^*$ and volume of large $|\Delta B0|$ have no clinical basis. Choosing different cut-off values may change the volumes. However, in a range of half the cutoff value up to two times the cutoff value, both for $T2^*$ and $|\Delta B0|$ maximum 1 shift was observed in the ranking of the markers. Also, in order to estimate the maximum effect of a marker on $T2^*$ and $\Delta B0$, we chose to use three and six voxels to calculate $frac_{\Delta T2^*, \max}$ and $\Delta_{\max} B0$, respectively. Selecting more voxels will decrease the values of $\Delta B0$ and $frac_{\Delta T2^*, \max}$. However, in our data, increasing the number of voxels selected to 6 and 12 only decreased the

values of $\text{frac}_{\Delta T^*, \max}$ and $\Delta_{\max} B_0$ on average by 6% and 13%, respectively. In addition, a maximum shift of 1 rank was observed in the ranking of the markers.

Even though T_2^* has been used with the assumption that it is independent of MR-settings, there are some restrictions to this assumption. The phantom findings were obtained at a $T_2^*_{\text{background}}$ of 43.3–47.0 ms. It is unclear how our T_2^* -maps in the presence of markers translate to regions where the initial T_2^* is different. However, pancreatic tissue has a T_2^* value of 41.5 ms (standard deviation 7.4 ms) and therefore, our data are well applicable to the pancreas [31]. In addition, T_2^* depends on the changes in B_0 within the voxel. When using markers, B_0 changes over distances comparable to the voxel size (Fig. 4.1), and T_2^* therefore depends on voxel size. However, we expect this effect to be similar for all markers and therefore, the ranking of the volume of low T_2^* and the ranking of $\text{frac}_{\Delta T^*, \max}$ will not change.

The correlation between the potential visibility and the potential artifacts is explained by the fact that both are produced by local changes in B_0 . A non-gold fiducial marker, especially manufactured for CT, CBCT, and MRI purposes, such as hydrogel [32], may offer visibility through a different mechanism and overcome this correlation.

Visibility and artifact sizes depend on the sequence parameters and can therefore be optimized by changing sequence parameters. For example, according to Eq. 4.3 lower BWs correspond to larger signal shifts and thus larger artifacts. The BW can thus be increased to obtain DW-images with smaller artifacts. However, increase in BW is limited when using single shot EPI read-outs, as done in most clinical DWI. Therefore, other ways of increasing the BW can be used, such as the multi-shot EPI or diffusion prepared non EPI sequences [33, 34]. However, such sequences are challenging to implement in the abdominal region. Also, sequences and reconstruction methods have been developed that specialize on visualizing and locating markers, such as the 3D center-out radial sampling with off-resonance reception imaging technique [35]. Likewise, these sequences are challenging to implement in the abdominal region. Ultimately, the optimal method for visualizing the markers and reducing the size of the artifact they cause will depend on the application. In the end, all the methods described above are based on the fact that the markers modify T_2^* and B_0 . We, therefore, focussed on the potential of a marker to be visible and to cause artifacts irrespective of the MR-sequence.

With exception of the markers containing iron, visibility and artifacts increase with marker size. Visibility and artifacts occur due to the differences in the magnetic susceptibility, χ , of gold ($\chi = -34 \times 10^{-6}$) and tissue (χ ranges from -11×10^{-6} to -7×10^{-6}), which are both diamagnetic [17]. Larger markers have more gold and thus have a larger effect on the magnetic field, making them better visible and creating

more artifacts. These findings decide with the findings of Chan et al. and the fact that markers with $\varnothing > 1$ mm can be detected on MR-images [11-14].

However, the markers that contained iron behaved differently; they were the smallest markers in this study, but had the best visibility and the largest artifacts. Iron is ferromagnetic ($\chi = 2 \times 10^5$) and therefore, iron affects B_0 more than gold [17]. It is interesting to note that, in our data, the markers containing iron produced a region of low B_0 surrounded by high B_0 , whereas this was the opposite for the non-iron markers (Fig. 4.1).

Therefore, if visibility is desired, we recommend using marker 2, for which we show *in vivo* visibility (Fig. 4.6). If other markers, with lower volumes of T_2^* and lower $\text{frac}_{\Delta T_2^*, \max}$ are chosen, they may become indiscernible *in vivo*. At what value of T_2^* and $\text{frac}_{\Delta T_2^*, \max}$ markers become indiscernible, ultimately depends on the sequence parameter settings, as discussed above.

When using DWI_{low BW}, marker 2 induced an artifact with a diameter of ~20 mm in the phantom. Changing to DWI_{high BW} only reduced the artifact extent to a diameter of ~15 mm. Usually, three markers are implanted in pancreatic tumors. Considering the size of these tumors (31 mm on average) [4], a fair amount of the tumor may contain unreliable information in DW-images.

The artifacts due to the markers were not distinguishable from anatomical features in the patient DW-images (Fig. 4.6). However, the phantom study showed that marker 2 disrupted the magnetic field in the surrounding volume (Fig. 4.1) which caused artifacts in the form of signal shifts in the DW-images (Fig. 4.5). We therefore believe that the *in vivo* data also contain shifted signal in a similar sized volume around a marker.

For the reasons mentioned above, we believe that if DWI is desired close to markers, a marker must be selected that gives minimal changes in B_0 . We recommend selecting one of the markers with low maximum shifts and small volume of affected B_0 , such as marker 3 or 5. For a typical clinical DWI measurement ($BW = 17.5$ Hz, 2×2 mm² voxels), markers 3 and 5 show shifts < 0.7 mm and show no abnormalities in our phantom measurements (Fig. 4.5). Therefore, we believe that DWI data close to these markers can be trusted (keeping in mind that data may be shifted up to 0.7 mm). Unfortunately, these markers were not visible in MR-images. When a different marker is selected, larger deformations of up to 9.8 mm for marker 2 may occur close to the marker in DW-images, preventing accurate tumor delineation.

As a result of this work, it is now possible to either select markers that are visible on MRI and allow for MRI-CT/MRI-CBCT image alignment, or to select markers that allow for DWI, which can improve tumour delineation, but lack visibility. The visible markers can lead to better patient alignment [1] whereas the non-artifact markers

lead to better tumour delineation [9, 10]. In both cases, the markers can potentially help to prevent unnecessary dose to surrounding organs at risk or an underdosage of the target. However, the improvement in clinical outcome has not been quantified yet.

Conclusion

Changes in $T2^*$ and $\Delta B0$ are sequence-independent measures for visibility and artifact size of markers in MRI. There was no ideal marker in our sample that is visible without showing artifacts on DW-images. Therefore, when DWI or other forms of artifact sensitive sequences are desired, ironless markers with small diameter and in an extended configuration are preferable (markers 3 or 5). When marker detection is desired, markers that contain iron and in a folded configuration (marker 2) are preferred, but this choice can make DWI unreliable close to the marker.

References

1. van der Horst A, Wognum S, Dávila Fajardo R et al. Interfractional position variation of pancreatic tumors quantified using intratumoral fiducial markers and daily cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 87(1):202–208.
2. Lens E, van der Horst A, Kroon PS et al. Differences in respiratory-induced pancreatic tumor motion between 4D treatment planning CT and daily cone beam CT, measured using intratumoral fiducials. *Acta Oncol.* 2014; 53(9):1257–1264.
3. Varadarajulu S, Bang JY, Hebert-Magee S. Assessment of the technical performance of the flexible 19-gauge EUS-FNA needle. *Gastrointest. Endosc.* 2012; 76(2):336–343.
4. Dávila Fajardo R, Lekkerkerker SJ, van der Horst A et al. EUS-guided fiducial markers placement with a 22-gauge needle for image-guided radiation therapy in pancreatic cancer. *Gastrointest. Endosc.* 2014; 79(5):851–855.
5. Yamazaki H, Nishiyama K, Tanaka E et al. Dummy run for a phase II multi-institute trial of chemoradiotherapy for unresectable pancreatic cancer: Inter-observer variance in contour delineation. *Anticancer Res.* 2007; 27(4 C):2965–2971.
6. Rasch C, Keus R, Pameijer FA et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 1997; 39(4):841–848.
7. Villeirs GM, van Vaerenbergh K, Vakaet L et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther. Onkol.* 2005; 181(7):424–430.
8. Mitchell DG, Snyder B, Coakley F et al. Early invasive cervical cancer: Tumor delineation by magnetic resonance imaging, computed tomography, and clinical examination, verified by pathologic results, in the ACRIN 6651/GOG 183 intergroup study. *J. Clin. Oncol.* 2006; 24(36):5687–5694.
9. Schima W. MRI of the pancreas: tumours and tumour-simulating processes. *Cancer Imaging* 2006; 6:199–203.
10. Matsuki M, Inada Y, Nakai G et al. Diffusion-weighted MR imaging of pancreatic carcinoma. *Abdom. Imaging* 2007; 32(4):481–483.
11. Huisman HJ, Fütterer JJ, van Lin ENJT et al. Prostate cancer: Precision of integrating functional MR imaging with radiation therapy treatment by using fiducial gold markers. *Radiology* 2005; 236(1):311–317.
12. Kapanen M, Collan J, Beule A et al. Commissioning of MRI-only based treatment planning procedure for external beam radiotherapy of prostate. *Magn. Reson. Med.* 2013; 70(1):127–135.
13. Nichol AM, Brock KK, Lockwood GA et al. A magnetic resonance imaging study of prostate deformation relative to implanted gold fiducial markers. *Int. J. Radiat. Oncol. Biol. Phys.* 2007; 67(1):48–56.
14. Chan MF, Cohen G ad N, Deasy JO. Qualitative evaluation of fiducial markers for radiotherapy imaging. *Technol. Cancer Res. Treat.* 2014; 14(3):1–7.
15. Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. *J. Magn. Reson. Imaging* 2006; 24(3):478–488.
16. Rylander S, Thörnqvist S, Haack S et al. Intensity profile based measurement of prostate gold markers influence on 1.5 and 3T diffusion-weighted MR images. *Acta Oncol.* 2011; 50(6):866–872.
17. Schenck JF. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Med. Phys.* 1996; 23(6):815–850.
18. Koch KM, Hargreaves BA, Pauly KB et al. Magnetic resonance imaging near metal implants. *J. Magn. Reson. Imaging* 2010; 32(4):773–787.
19. Hargreaves BA, Worters PW, Pauly KB et al. Metal-induced artifacts in MRI. *Am. J. Roentgenol.* 2011; 197(3):547–555.

20. Wersall P, Hartwig A, Castellanos E et al. Spatial Accuracy and Visibility Studies of a MRI Compatible Fiducial Marker [ASTRO abstract 3382]. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 84(Suppl.):3:S737–S738.
21. Yoshimura K, Kato H, Kuroda M et al. Development of a tissue-equivalent MRI phantom using carrageenan gel. *Magn. Reson. Med.* 2003; 50(5):1011–1017.
22. Kato H, Kuroda M, Yoshimura K et al. Composition of MRI phantom equivalent to human tissues. *Med. Phys.* 2005; 32(10):3199–3208.
23. Herráez MA, Burton DR, Lalor MJ, Gdeisat MA. Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path. *Appl. Opt.* 2002; 41(35):7437–7444.
24. Froeling M, Nederveen AJ, Heijtel DFR et al. Diffusion-tensor MRI reveals the complex muscle architecture of the human forearm. *J. Magn. Reson. Imaging* 2012; 36(1):237–248.
25. Graf H, Steidle G, Martirosian P et al. Metal artifacts caused by gradient switching. *Magn. Reson. Med.* 2005; 54(1):231–234.
26. Lauer UA, Berger A, Schick F. RF artifacts caused by metallic implants or instruments which get more prominent at 3 T : An in vitro study. *Magn. Reson. Imaging* 2005; 23(3):493–499.
27. Graf H, Steidle G, Martirosian P et al. Effects on MRI due to altered rf polarization near conductive implants or instruments. *Med. Phys.* 2006; 33(1):124–127.
28. Vashaee S, Goora F, Britton MM et al. Mapping B1-induced eddy current effects near metallic structures in MR images: A comparison of simulation and experiment. *J. Magn. Reson.* 2015; 250:17–24.
29. Riffel P, Rao RK, Haneder S et al. Impact of field strength and RF excitation on abdominal diffusion-weighted magnetic resonance imaging. *World J. Radiol.* 2013; 5(9):334–344.
30. Sandrasegaran K, Lin C, Akisik FM, Tann M. State-of-the-art pancreatic MRI. *Am. J. Roentgenol.* 2010; 195(1):42–53.
31. Schwenzer NF, Machann J, Haap M et al. T2* Relaxometry in Liver, Pancreas, and Spleen in a Healthy Cohort of One Hundred Twenty-Nine Subjects-Correlation With Age, Gender, and Serum Ferritin. *Invest. Radiol.* 2008; 43(12):854–860.
32. de Souza Lawrence L, Ford E, Gilbert C et al. Novel applications of an injectable radiopaque hydrogel tissue marker for management of thoracic malignancies. *Chest* 2013; 143(6):1635–1641.
33. Cheung MM, Wu EX. Diffusion imaging with balanced steady state free precession. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, 2012; 2012(852):90–93.
34. Skare S, Newbould RD, Clayton DB et al. Clinical multishot DW-EPI through parallel imaging with considerations of susceptibility, motion, and noise. *Magn. Reson. Med.* 2007; 57(5):881–890.
35. Seevinck PR, de Leeuw H, Bos C, Bakker CJG. Highly localized positive contrast of small paramagnetic objects using 3D center-out radial sampling with off-resonance reception. *Magn. Reson. Med.* 2011; 65(1):146–156.

Supplementary Materials

Figure 4.A. shows the data from both read-out directions and both orientations of the marker, which is the data that Fig. 4.4 is based on.

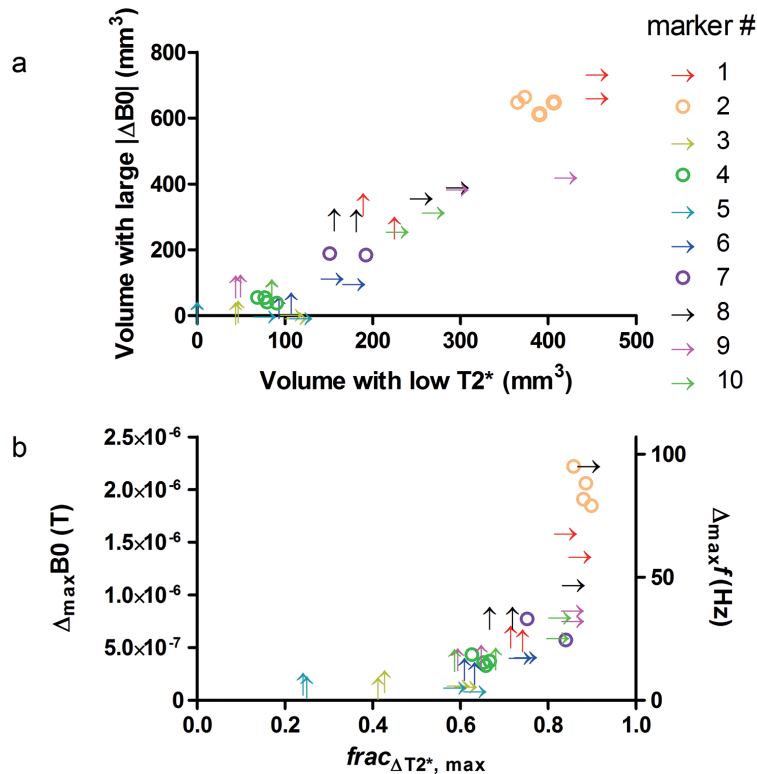


Figure 4.A. (a): Plot of the volume with signal shifts plotted against the volume with signal voids. (b): Plot of magnitude of the signal shifts against the magnitude of the signal voids. Arrows pointing upward indicate markers parallel to the B_0 field, arrows pointing to the right indicate markers perpendicular to B_0 and circles indicate markers with a folded configuration. Note that for marker 7 the data obtained with the FH read-out was excluded from the analyses, as described in the Results.

Table 4.A shows the calculated maximum distances voxels could theoretically shift with respect to each other, given the measured $\Delta_{\max} B_0$, for both diffusion protocols.

Table 4.A. The calculated maximum shifts for the ten investigated markers.

Marker #	Shift DWI _{low BW} [mm]	Shift DWI _{high BW} [mm]
1	5.0	2.2
2	9.8	4.4
3	0.7	0.3
4	1.8	0.8
5	0.6	0.3
6	1.7	0.7
7	3.3	1.5
8	3.0	1.3
9	6.0	2.7
10	2.6	1.2



CHAPTER 5

Artifacts of biliary stents

Quantitative assessment of biliary stent artifacts on MR images: Potential implications for target delineation in radiotherapy

Oliver J Gurney-Champion
Thijs Bruins Slot
Eelco Lens
Astrid van der Horst
Remy Klaassen
Hanneke WM van Laarhoven
Geertjan van Tienhoven
Jeanin E van Hooft
Aart J Nederveen
Arjan Bel

Medical Physics 2016;43(10):5603–5615

Abstract

Purpose: Biliary stents may cause susceptibility artifacts, gradient-induced artifacts and radio frequency (RF) induced artifacts on magnetic resonance images, which can hinder accurate target volume delineation in radiotherapy. In this study, we investigated and quantified the magnitude of these artifacts for stents of different materials.

Methods: Eight biliary stents made of nitinol, platinum-cored nitinol, stainless steel or polyethylene from seven vendors, with different lengths (57–98 mm) and diameters (3.0–11.7 mm) were placed in a phantom. To quantify the susceptibility artifacts sequence-independently, ΔB_0 -maps and T_2^* -maps were acquired at 1.5 and 3 T. To study the effect of the gradient-induced artifacts at 3 T, signal decay in images obtained with maximum readout gradient-induced artifacts was compared to signal decay in reference scans. To quantify the RF induced artifacts at 3 T, B_1 -maps were acquired. Finally, ΔB_0 -maps and T_2^* -maps were acquired at 3 T of two pancreatic cancer patients who had received platinum-cored nitinol biliary stents.

Results: Outside the stent, susceptibility artifacts dominated the other artifacts. The stainless steel stent produced the largest susceptibility artifacts. The other stents caused decreased T_2^* up to 5.1 mm (1.5 T) and 8.5 mm (3 T) from the edge of the stent. For sequences with a higher bandwidth per voxel (1.5 T: $BW_{\text{vox}} > 275$ Hz/voxel; 3 T: $BW_{\text{vox}} > 500$ Hz/voxel), the B_0 -related susceptibility artifacts were negligible (< 0.2 voxels). The polyethylene stent showed no artifacts. *In vivo*, the changes in B_0 and T_2^* induced by the stent were larger than typical variations in B_0 and T_2^* induced by anatomy when the stent was at an angle of 30° with the main magnetic field.

Conclusions: Susceptibility artifacts were dominating over the other artifacts. The magnitudes of the susceptibility artifacts were determined sequence-independently. This method allows to include additional safety margins that ensure target irradiation.

Introduction

Magnetic resonance imaging (MRI) is increasingly used for radiotherapy planning as, for some tumor sites, it has been shown to improve the accuracy of tumor delineation due to its high soft tissue contrast compared to CT [1-3]. Also, considering recent developments in magnetic resonance (MR) guided radiotherapy systems, MRI can now be used for daily patient position verification and monitoring of tumor position during irradiation [4, 5]. In recent years, diffusion-weighted MRI (DWI) is gaining attention as it enables tumor detection, tumor characterization, and treatment response monitoring [6-10]. However, a major drawback of DWI is that the images are prone to artifacts [11].

Radiation treatment planning of pancreatic cancer patients may especially benefit from the increased tumor visibility and healthy tissue contrast on MRI as tumor delineation on CT, which is the current standard, is extremely challenging [12-14]. However, many pancreatic cancer patients receive implants that may cause artifacts on MR-images. For example, at our institute, all pancreatic cancer patients scheduled for radiotherapy receive intratumoral gold fiducials to guide and improve cone-beam CT-based patient setup [15-17]. These fiducials can cause considerable artifacts, depending on the specific fiducial and MRI settings used [18]. In addition, 50%–70% of these patients suffer from jaundice at the time of diagnosis [19, 20]. To relieve jaundice, these patients receive a biliary stent before the start of radiotherapy. These stents are known to cause artifacts [21, 22], such as susceptibility artifacts [23], gradient-induced artifacts [24, 25] and radio frequency (RF)-induced artifacts (also known as RF-shielding) [26-29]. As the stent is placed in the vicinity of the tumor [30], MRI artifacts caused by these stents may hinder tumor delineation.

For radiotherapy treatment planning, accurate target delineation is required. As the accuracy of delineation may be affected by imaging artifacts, one would like to know the potential magnitude and size of artifacts induced by a stent. Uncertainties in tumor extent caused by artifacts can then be taken into account as an additional factor among the other uncertainties in the definition of treatment field safety margins, to ensure irradiation of the target. Hence, uncertainties in delineation due to MRI artifacts must be quantified as a function of distance from the stent. Such a measure is not described in the current literature.

Therefore, the goal of this study was to investigate and quantify the severity of image artifacts as a function of distance from the stent for any sequence, for eight different clinically used biliary stents.

Materials and Methods

MR images were acquired with an Ingenia 3 T scanner (Philips Healthcare, Best, the Netherlands), using a 16-channel phased array anterior coil and a 10-channel phased array posterior coil, as well as with an Achieva 1.5 T scanner (Philips Healthcare, Best, the Netherlands), using an 8-channel phased array anterior coil and an 8-channel phased array posterior coil. Since artifact size and magnitude grow with magnetic field strength, the most prominent artifacts, the susceptibility artifacts, were also studied at 1.5 T.

As the orientation of the stents relative to the magnetic field influences the behavior of MRI artifacts, all phantom measurements were repeated with the stents in two different orientations: one orientation with the central axis parallel to the main magnetic field (within 5°) and one with the central axis perpendicular to the main magnetic field (within 5°). From here onwards, these orientations will be referred to as parallel orientation and perpendicular orientation, respectively (Fig. 5.1). A researcher aligned the stent manually and checked the alignment on a high bandwidth (1602 Hz/voxel, thus negligible deformation) steady-state free precession (SSFP) image with Dixon reconstruction. This procedure was repeated iteratively until the stent was aligned at a 5° accuracy. Note that *in vivo*, the orientation of the stent cannot be changed.

Unless mentioned otherwise, data were analyzed in MATLAB 2014b (Math-Works, Natick, USA). Plots were produced in Prism 5 (GraphPad Software, La Jolla, CA). MR images used in figures were exported from MATLAB and ITK-snap 2.2.0 [31].

We investigated three types of artifacts: susceptibility artifacts, gradient-induced artifacts and RF-induced artifacts (Table 5.1), which are discussed in detail in Sec. I of the supplementary materials [32].

Phantom study

We selected eight stents consisting of various materials and from seven vendors (Table 5.2, Fig. 5.1). To guide stent placement, five of these stents contained radiopaque markers at their tips that are clearly visible on a radiograph (Fig. 5.1). The magnetic susceptibilities (χ , also known as volume susceptibility) of the stent materials were $\chi = 245 \times 10^{-6}$ (nitinol), $\chi = 279 \times 10^{-6}$ (platinum), $\chi = 3520\text{--}6700 \times 10^{-6}$ (stainless steel) and $\chi = 1.86 \times 10^{-6}$ (polyethylene) [23, 33]. Susceptibility artifacts are expected to be more pronounced for larger differences in χ between stent and tissue ($\chi = -11 \cdot 10^{-6}$ to $-7 \cdot 10^{-6}$) [23]. Some stents were covered with a silicone layer. Silicone has $\chi = -4.2 \times 10^{-6}$ and the effect on susceptibility from the silicone layer is thus inferior to the effects from the metal stent [23]. The electrical resistivities were $\rho = 76 \times 10^{-6} \Omega \text{ cm}$

(nitinol), $\rho = 11 \times 10^{-6} \Omega \text{ cm}$ (platinum), $\rho = 72 \times 10^{-6} \Omega \text{ cm}$ (stainless steel) and $\rho > 1 \times 10^{15} \Omega \text{ cm}$ (polyethylene) [34–37]. Stents from more conducting materials are expected to cause larger gradient-induced and RF-induced artifacts.

Agar is often used as phantom material [38–40] as it allows studying isolated effects from implants in a gel with tissue-like properties [41, 42]. Therefore, we prepared eight jars (radius = 105 mm, height = 205 mm) filled with agar gel containing a fully expanded stent at their center. The agar gel mimicked tissue-like properties. We achieved this by mixing 1.3 g NaCl, 120 μmol Gadovist, 35 g agar and 1.5 g sodium benzoate, per liter demi-water [41, 42]. The agar gel had a $T_1 = 851 \text{ ms}$ (range 784–1065 ms) and $T_2 = 41 \text{ ms}$ (range 38–45 ms) at 3 T, which is in the range of liver and pancreatic tissue relaxation properties [43]. First, the jar was half filled with agar gel. Then, we used an extension clamp to hold the stent partially submerged while the gel was setting. Once set, the extension clamp was removed, the orientation of the stent was marked on the edge of the phantom, and the jar was filled with agar gel. We scanned one jar at a time. To reduce field inhomogeneities at the edge of the jar, the jar was placed on a holder at the center of a large water-filled container (40 \times 30 cm^2).

Table 5.1. The assessed artifacts with their source, effects and the performed measurement used to evaluate them.

Artifact	Source	Effects	Measurement
Susceptibility artifacts [23]	Difference in magnetic susceptibility between stent and its surroundings locally alter B_0 .	Signal shifts and hypointense regions.	ΔB_0 -maps and T_2^* -maps.
Gradient-induced artifact [24, 25]	The onset of the frequency encoding gradient induces an eddy current in the stent, which alters the local B_0 for a brief period.	Hypointense regions.	Two images of the stent positioned off-center acquired with a read-out direction that (1) induced eddy currents and (2) minimized eddy currents.
RF-induced artifact [26–28]	RF-pulses cause an eddy current in the stent that changes the B_1 field close to the stent.	Altered flip angle around the stent.	B_1 -maps.

Susceptibility artifacts

As we described earlier [18, 44], direct imaging of the local change in B_0 (ΔB_0) and T_2^* properties around the stent is a sequence-independent measure for the susceptibility artifacts. Therefore, we acquired a single eight echo T_1 -weighted SSFP image (Table 5.3) with the stent at the center of the bore. From the magnitude and phase data, we

Table 5.2. Stents tested.

	Manufacturer	Stent material	Marker material	Length (mm)	Diameter (center/tip, mm)	Covered stent ^a
1: X-suit NIR	Medinol (Tel Aviv, Israel)	Nitinol ^b	Tantalum	77	9.3/10.7	Yes
2: S.M.A.R.T. CONTROL	Cordis Corporation (Fremont, USA)	Nitinol ^b	Tantalum	62	9.2/9.7	No
3: Zilver 635 Biliary Stent	Cook Medical (Bloomington, USA)	Nitinol ^b	Gold	81	10.3	No
4: SX-ELLA Stent Biliary – Nitinella Plus	ELLA-CS (Hradec Kralove, Czech Republic)	Nitinol ^b	Platinum-iridium	82	9.5	No
5: WallFlex Biliary RX Stent	Boston Scientific (Marlborough, USA)	Platinum-cored nitinol ^c	n.a.	98	9.9/11.7	No
6: Evolution Biliary stent	Cook medical (Bloomington, USA)	Platinum-cored nitinol ^c	n.a.	62	10.0/10.9	Yes
7: Visi-Pro	Covidien (Plymouth, USA)	Stainless steel	Tantalum	57	8.5	No
8: PE-stent	ENDO-FLEX GmbH (Voerde, Germany)	Poly-ethylene	n.a.	71	3.0	n.a.

Note: The columns Length and Diameter give the measured length and diameter of the expanded stent. When the stents were broader at the tips than in the center (Fig. 5.1) both stent diameters are given.

^a Covered with a silicone layer.

^b Nitinol is an alloy of nickel and titanium.

^c Platinum-cored nitinol consists of a platinum core with a nitinol coating.

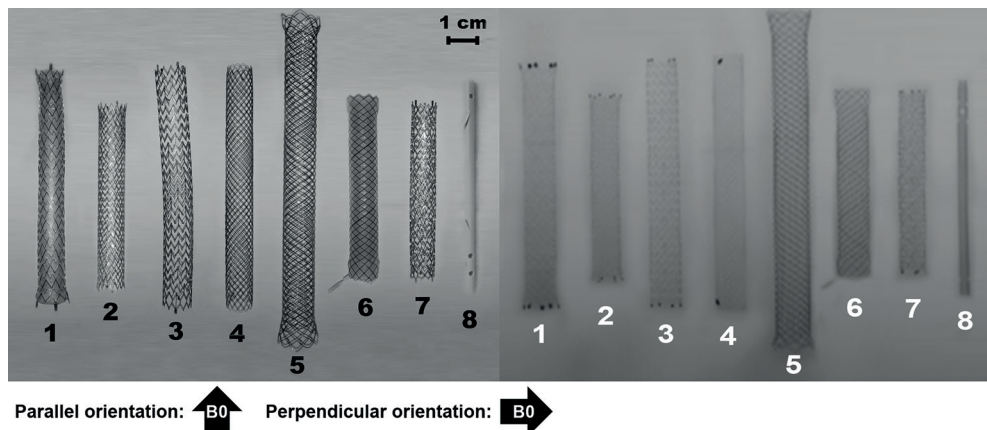


Figure 5.1. Composition photograph (**left**) and radiograph (**right**) of the stents. Weaving differed between stents. The radiopaque markers can be seen in the radiographs as dark spots at the tips of stents #1–4 and #7. The 120 kV radiographs were obtained using the flat-panel from an Elekta cone beam CT (Elekta Oncology systems, Crawley, UK). The different directions of B0 with respect to the stents are also illustrated.

produced a $T2^*$ -map and a $\Delta B0$ -map, showing per voxel local $T2^*$ and local change in $\Delta B0$, respectively. Maps were produced using MATLAB 2014b for the $T2^*$ -maps and DTITools for Mathematica [45], Mathematica 10.3 (Wolfram Research, Inc., Oxfordshire, UK) for the $\Delta B0$ -maps.

The $\Delta B0$ -maps contained effects from the stent as well as background variations in $\Delta B0$. To remove the low-frequency background variations, we subtracted a low-pass filtered version of the $\Delta B0$ -map from the original $\Delta B0$ -map. This technique is discussed in more detail in Sec II of the supplementary materials [32]. To improve the signal-to-noise ratio, we averaged the $\Delta B0$ or $T2^*$ values from a 9–60 mm broad region along the length of the stent (depending on the size of the region where the effects of the stent were homogeneous, mean 36 mm). This averaged value was plotted as a function of distance from the stent edge (Sec. II of the supplementary material [32] Fig. 5.B shows an example). Data from this region were especially interesting as the tumor is likely to be located here. The position of the edge of the stent was determined visually, under the guidance of the magnitude image of the SSFP and a reference Dixon scan and the diameter was compared to the measured diameter (Table 5.2) as quality control. The position of the edge of the stent was always set at a voxel location and, hence, the maximum error was ± 0.5 voxel. The standard deviation of $B0$ or $T2^*$ values from a homogeneous region was taken as an estimate of the precision.

As the susceptibility artifacts depend on field strength, we repeated the measurements, at 1.5 T, for which we used a similar sequence to the scans done at 3 T (Table 5.3).

To illustrate the effect of orientation on the susceptibility artifacts on a typical stent, we also acquired SSFP images with the stent's central axis at nine different angles (0° – 90°) on the main magnetic field for a typical stent (stent #6) at 3 T.

Gradient-induced artifacts

To assess the effect of the gradient-induced artifacts, we also acquired SSFP images with the stent 13.2 ± 2.0 (mean \pm standard deviation) cm off-center in the foot-head (FH) direction. In our SSFP sequence, the read-out gradient should induce the largest artifact, as its slew rate (96 mT/m/ms) and gradient strength (21 mT/m) were higher than for the slice-selecting gradient (slew rate = 25 mT/m/ms; gradient strength = 2.8 mT/m). The eddy currents from phase encoding, slice encoding and diffusion weighting (independent of b-value) gradients are in opposite direction for the ramp-up and ramp-down of the gradient and, hence, their effect is canceled out at the time of signal acquisition [24]. Therefore, acquisitions were done with two different frequency encoding directions: a reference scan with the encoding in the right-left (RL) direction, which does not allow for switching frequency encoding gradients to

Table 5.3. MRI sequence acquisition parameters.

	Phantom			Clinical protocol		
	3 T		1.5 T	3 T		
	SSFP (ΔB_0 and T_2^*)	DREAM (B_1)	SSFP (ΔB_0 and T_2^*)	Dixon	SSFP	DWI
Dimensions	3D	2D	3D	3D	3D	2D
FOV (mm ³)	407 × 407 × 94	300 × 246 × 148	407 × 407 × 94	400 × 350 × 95 (RL × AP × FH)	400 × 350 × 95 (RL × AP × FH)	430 × 110 × 70 (RL × AP × FH)
Resolution (mm ²)	1.7 × 1.7	2.7 × 2.7	1.7 × 1.7	1.7 × 1.7	2.3 × 2.3	3.0 × 3.0
Slice thickness (mm)	1.7	5 ^a	1.7	1.7	4.6	3.7
TR/TE/ ΔTE (ms)	20/2.3/2.3	5/1.6/0.7	25/4.6/4.6	4.7/1.15/1.0	20/2.3/2.3	4000/45/—
Parallel imaging	2	—	2	2/1.5 (AP/FH)	2/1.5 (AP/FH)	1.7
Half Fourier	—	—	—	—	—	0.8
Flip angle (°)	12	10	21	8	12	90
BW_{vox} (Hz/voxel)	1488	2565	1425	1602	1973	58.8
Scan time (s)	63	50	163	20	23	148
Respiratory compensation	—	—	—	Breath hold	Breath hold	Navigator triggered

Abbreviations: SSFP, steady state free precession; DWI, diffusion-weighted imaging; FOV, field of view; TR, repetition time; TE, echo time; ΔTE , increase in TE; RL, right-left; AP, anterior-posterior; FH, foot-head; BW_{vox} , bandwidth per voxel.

^a with 0.5 mm slice gap

produce eddy currents, and an acquisition with the frequency encoding gradient in the FH direction, which should maximize eddy currents in the stent. If eddy currents in the stent have any effect on the image, the signal intensity close to the stent is expected to decrease for the second acquisition compared to the reference scan.

To assess the additional contribution of eddy currents to the signal decay around the stent, we compared the signal decays that occurred around the stent for the acquisition that maximized eddy currents, to the signal decay in the reference scan. As MR-signals have no absolute units, signal decay was normalized to the background signal and thus defined as the ratio between the signal from a ROI around the stent and a ROI containing only background signal. The ratio of the signal intensities should be higher in the reference scan if the gradient-induced artifact had any meaningful contribution.

RF-induced artifacts

To assess the effect on $B1$ homogeneity due to the RF-induced artifact, we placed the stent at the center of the bore and obtained $B1$ -maps with the DREAM sequence (Table 5.3) [46]. The background effect, induced by the phantom, was filtered out of the $B1$ -map using a similar method as for the $\Delta B0$ -map. After filtering we added a flip angle of 100% to each voxel again, to study the impact of the stent on a perfect situation of flip angle = 100%. After filtering, similar plots were made as for $T2^*$ and $\Delta B0$, taking the mean value from a 14–59 mm broad region (mean 37 mm) along the length of the stent.

Clinical sequences

To show the effect of the artifacts on clinical sequences, we acquired a DWI image, which is sensitive to artifacts. The DWI images were acquired at $b = 0$ and $600 \text{ mm}^2\text{s}$. We also acquired images using a sequence that is known to be more robust to susceptibility artifacts, the three echo $T1$ -weighted SSFP sequence. For this scan, we used the vendor supplied Dixon reconstruction without $T2^*$ correction (Table 5.3, clinical protocol). All images were obtained using the geometric distortion correction package provided by the vendor, which is on by default.

Patient study

To assess stents in a more clinical situation in which stents may be deformed and contain air, we selected six patients that received metallic biliary stents as part of our standard clinical procedure and from whom MRI data were obtained. All patients had given written informed consent for the additional MRI. They were scanned as part of ongoing studies (NCT01989000, NCT01995240), which were approved by the ethics committee of our institute.

Patients were positioned in the scanner according to standard clinical protocol. Therefore, the stent orientation and position were typical for clinical situations, meaning that the stent was not necessarily aligned with the main magnetic field or at the center of the MRI bore. Due to limited scan time, we only acquired the multiecho SSFP images as they allowed us to investigate the most dominant artifacts. The protocols were adapted slightly from the phantom measurements to deal with respiratory motion (Table 5.3). Also, we obtained standard clinical images to illustrate the effect of stents *in vivo* in two of these patients (patient 1 and 2). Patient 1 received stent #6 and patient 2 received stent #5.

To best represent the possible *in vivo* variations, we selected four patients (patients 3–6) in which the angle between the stent and the main magnetic field varied

between 0° (patient 3), 25° (patient 4), 35° (patient 5), and 45° (patient 6). Patient 6 had air inside the stent. Patient 3 had a stent which was bent throughout the duct, but the artifacts caused by the stent were assessed at a location where it had an angle of 0° . Patients 4–6 had a large region over which the stent remained straight, making a comparison with the phantom more straightforward than for patient 3. Patient 3–5 received stent #6, whereas patient 6 received stent #1.

The magnitude image of the initial echo of the SSFP-sequence was used to determine the angle of the stent with the main magnetic field. Using imageJ [47], the images were first resliced such that slices were perpendicular to the stent's central axis. Then, ΔB_0 was determined in two perpendicular directions within the slice in a similar fashion as in the angle-dependent phantom experiment: the direction parallel to the projection of B_0 on the slice, and the direction perpendicular to the projection of B_0 on the slice. To minimize anatomical effects, ΔB_0 was determined over multiple (4–8, mean 6) slices and averaged. Patient anatomy caused a poor homogeneity in the background ΔB_0 that did not allow for smoothing, as was done in the phantom data. Instead, ΔB_0 plots were normalized by subtracting the mean ΔB_0 value from 4 to 6 data points along the plot that were far enough (range: 0.9–2.5 cm) from the stent that no effects were visually observed.

Results

Phantom study

Susceptibility artifacts

The obtained ΔB_0 -maps and T_2^* -maps of stents #6 (platinum-cored nitinol) and #7 (stainless steel) are shown in Fig. 5.2. Stent #6 is shown as an example as it was used *in vivo* while stent #7 is shown because it produced the largest artifacts. The maps after post-processing are shown for all stents in Sec. III of the supplementary material [32]. From these maps, it was clear that the artifacts caused by the stents were dominating over the artifacts caused by the radiopaque markers, as no additional contributions from these markers were visible. Therefore, no quantitative analysis was done on the radiopaque markers.

Except for the polyethylene stent (stent #8), all stents disrupted B_0 and T_2^* to some extent (Fig. 5.3). The size and magnitude of the disruptions corresponded to the χ of the stent materials, in which higher χ resulted in larger disruptions in size and magnitude. The stainless steel stent (stent #7) showed the largest disruption. The stents that were made from platinum-cored nitinol (stent #5 and #6) showed the

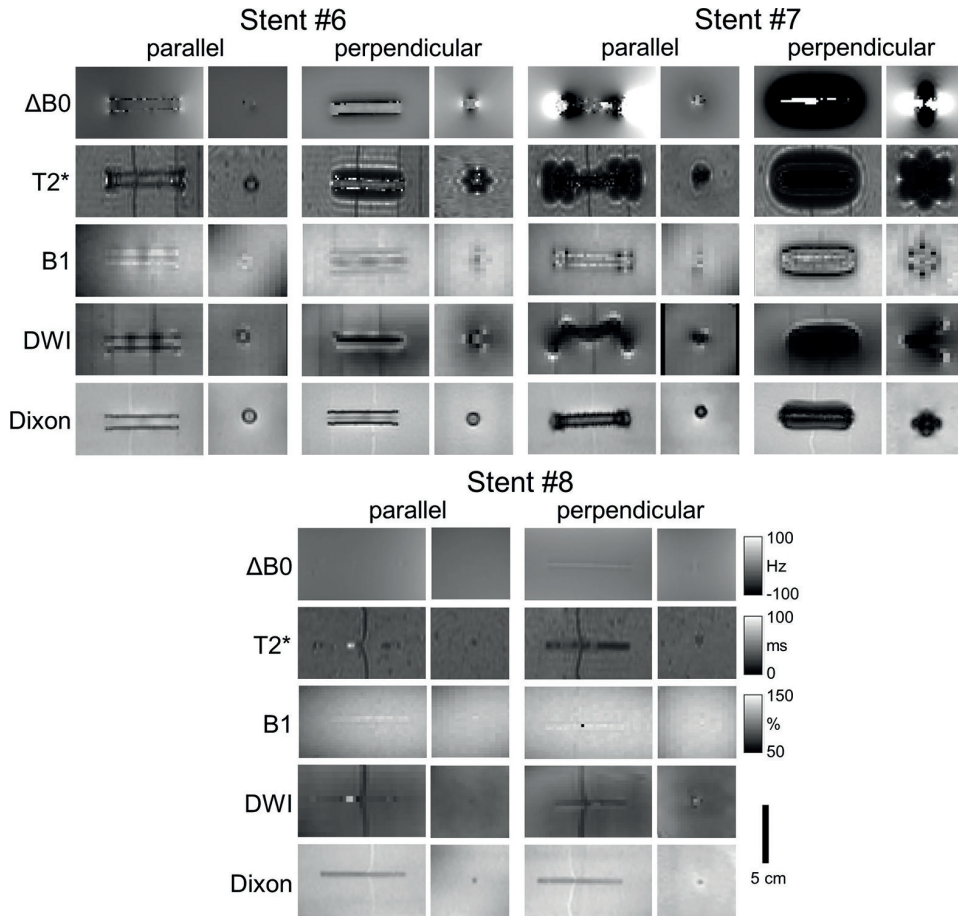


Figure 5.2. ΔB_0 -maps (top row), T_2^* -maps (second row), B1-maps (third row), DWI ($b = 0 \text{ mm}^2/\text{s}$, fourth row) and Dixon (bottom row) images of the stent used *in vivo* (stent #6, left), the stent that showed the largest artifacts (stent #7, middle) and the stent that caused the smallest artifact (stent #6, right). Images are shown for the aligned and perpendicular orientation, with two views per orientation. The vertical lines, in particular visible in the images of stent #8, are a boundary between two agar layers. Maps after post-processing are shown in Sec. III of the supplementary material [32].

second largest disruptions but had similar artifact sizes as the nitinol stents (stent #1–4). Finally, the polyethylene stent (stent #8) showed no disruptions. Note that the apparent increase of T_2^* to values higher than the agar gel ($\pm 45 \text{ ms}$), was the result of Gibbs ringing [48]. These artifacts are inherent to the artificial homogeneity of the agar gel. Also, the first data point (distance from stent = 0 mm) contained the stent's edge, and therefore, can be incorrect. The susceptibility artifacts were larger at 3 T than at 1.5 T (Figs. 5.3 and 5.4).

$T2^*$ and $\Delta B0$ strongly depended on stent orientation (Fig. 5.3) and increased gradually with angle between stent and main magnetic field (Fig. 5.5). $T2^*$ and $B0$ were affected most at the edge of the stent when the stent was oriented perpendicular to the main magnetic field whereas $T2^*$ and $B0$ were affected most at the tips of the stent when the stent was oriented parallel to the main magnetic field (Fig. 5.2)

Gradient-induced artifacts

The signal decay was similar for the images with eddy currents as for the reference images (Fig. 5.6). Thus, the gradient-induced artifacts were negligible compared to the susceptibility artifacts.

RF-induced artifacts

Except for stent #7, the effect of the stent on $B1$ was fairly local and $B1$ was 80%–110% of the desired $B1$ outside the stent (Fig. 5.7). We focus on $B1$ as a function of distance from the stent, which is not representative for the inside of the stent. Inside the stent shielding occurred, which affected $B1$ more severely (Sec. IV of the supplementary material [32]).

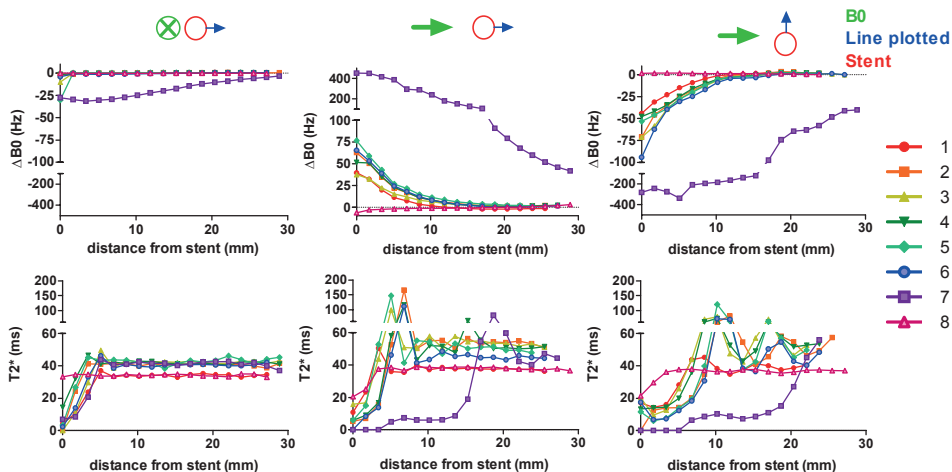


Figure 5.3. Plots of $\Delta B0$ (top) and $T2^*$ (bottom) as a function of distance from the stent edge for all stents at 3 T. Results are shown for three stent orientations: aligned orientation with plotted line perpendicular to the stent (left); perpendicular orientation, plotted line parallel to the main magnetic field (middle); and perpendicular orientation with plotted line perpendicular to the main magnetic field (right). Note that the axes are split. The maximum error along the x-axis was 0.9 mm, whereas the precision along the y-axis was 2.1 Hz (range between stents 0.9–3) and 2.4 ms (range 1.3–4). Additional plots are shown in Sec. IV of the supplementary material [32].

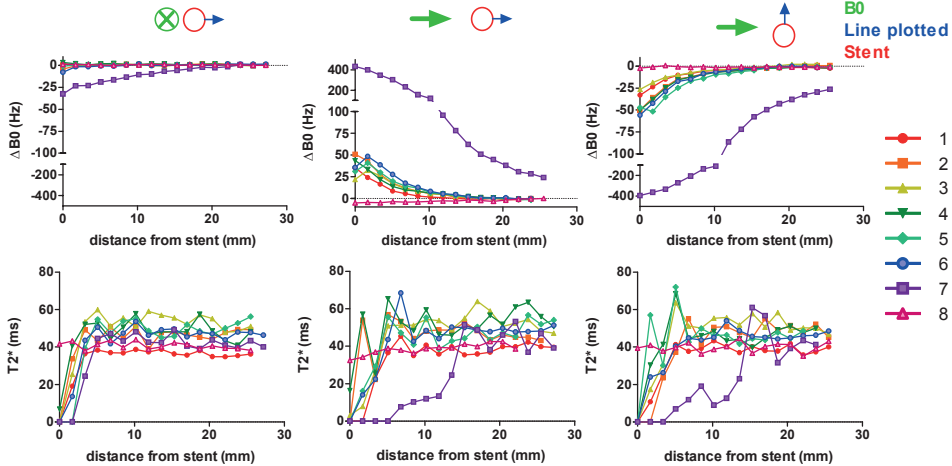


Figure 5.4. Plots of ΔB_0 (top) and T_2^* (bottom) as a function of distance from the stent edge for all stents at 1.5 T. Results are shown for three stent orientations: aligned orientation with plotted line perpendicular to the stent (left); perpendicular orientation, plotted line parallel to the main magnetic field (middle); and perpendicular orientation with plotted line perpendicular to the main magnetic field (right). Note that the axes are split. The maximum error along the x-axis was 0.9 mm, whereas the precision along the y-axis was 2.0 Hz (range between stents 1.6–3.5) and 7.8 ms (range 5.1–10). Additional plots are shown in Sec. IV of the supplementary material [32].

Clinical sequences

The discussed artifacts were larger in the DWI images than in the Dixon images (Fig. 5.2). Also, larger disruptions in the ΔB_0 -map, T_2^* -map, and B_1 -map (stent #7 compared to stent #6) corresponded to larger artifacts in the clinical image.

Patient study

In vivo, the surrounding tissue was less homogeneous compared to the phantom (Fig. 5.8), which caused local disturbances in B_0 unrelated to the stent. Also, due to the curvature of the biliary duct, stents were deformed *in vivo*. In vivo, stents #5 and #6 showed no large artifacts in the Dixon image (Fig. 5.8). Due to the curvature of the stent in patient 1, it was possible to show images at a location where the stent was parallel to the main magnetic field and at a location where the stent at an angle of 30° to the main magnetic field. Upon visual inspection, we noted that the part of the stent with an angle of 30° to the main magnetic field showed changes in ΔB_0 that were larger than normal variations in ΔB_0 throughout the patient caused by differences in susceptibility between tissues. The parts of the stents with an angle of 0° to the main

magnetic field showed changes in $\Delta B0$ that were smaller than normal variations in $\Delta B0$ throughout the patient caused by differences in susceptibility between tissues.

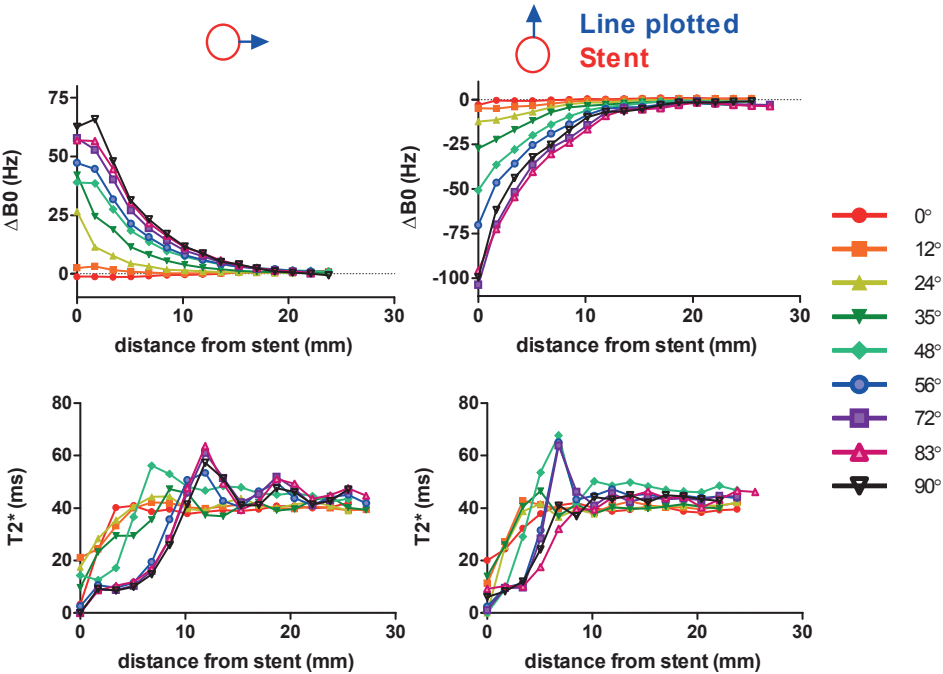


Figure 5.5. Plot of $\Delta B0$ (top) and $T2^*$ (bottom) as a function of distance from stent #6 at 3 T. Lines are plotted for different angles (0°–90°) between stent and main magnetic field. The maximum error along the x-axis was 0.9 mm, whereas the precision along the y-axis was 1.4 Hz (range between stents 0.9–1.9) and 2.0 ms (range 1.6–2.8).

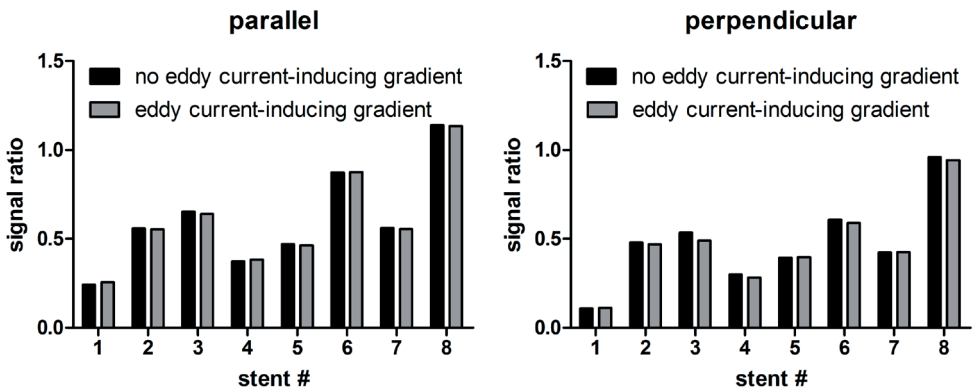


Figure 5.6. Graph of signal ratios from the ROI containing the stent and the ROI containing background signal for the acquisition that maximized readout gradient-induced eddy currents and the reference scan.

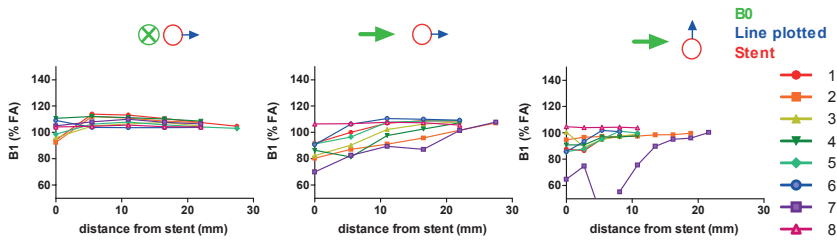


Figure 5.7. Plots of B_1 as a function of distance from the stent edge for all stents at 3 T. Results are shown for three stent orientations: aligned orientation with plotted line perpendicular to the stent (**left**), perpendicular orientation, plotted line parallel to the main magnetic field (**middle**) and perpendicular orientation with plotted line perpendicular to the main magnetic field (**right**). The maximum error along the x-axis was 1.4 mm in plane and 5.5 in the slice direction, whereas the precision along the y-axis was 2.2% (range between stents 1.3–4.1). The analysis was only performed up to a distance at which the B_1 was affected by the stent. Additional plots are shown in Sec IV of the supplementary material [32].

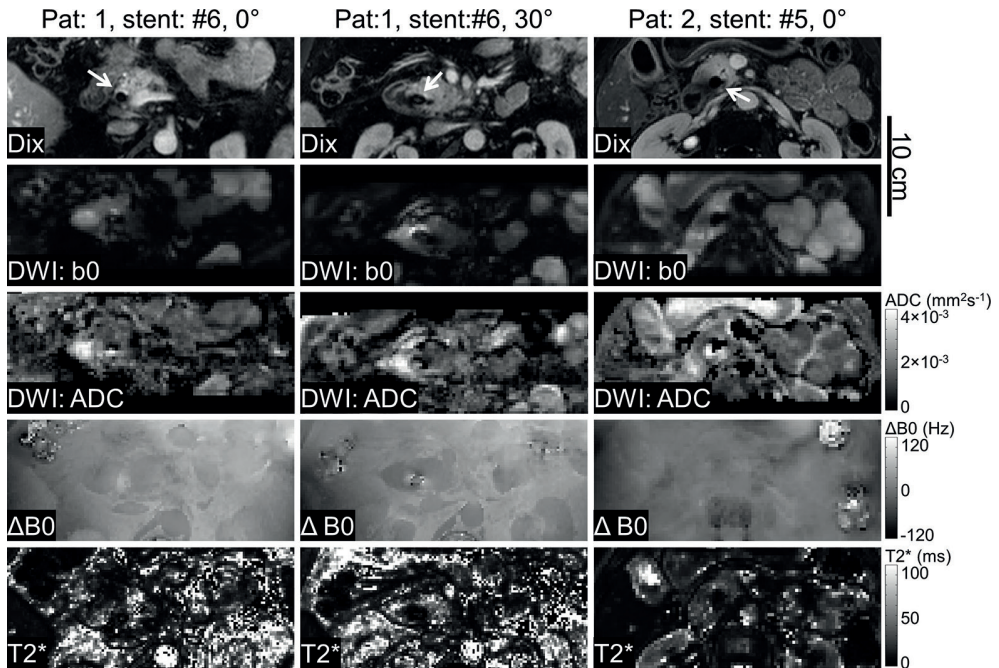


Figure 5.8. A typical example of Dixon images, DWI images ($b = 0 \text{ mm}^2/\text{s}$), the ADC image, ΔB_0 -maps and T_2^* -maps of stents #5 and #6 in two patients with pancreatic cancer. The selected slices show the stents under 0° and (for patient 1) 30° on the main magnetic field. The displayed slices from patient 1 were 15.3 mm apart. The white arrows indicate the stent location.

The *in vivo* measurements agreed reasonably well with the phantom measurements (Fig. 5.9). Note that the stent in patient 3 was parallel with B_0 in the central slice we took the data from but was bending from 24° in left–right orientation to 34° in anterior–posterior direction in the slices above and below. The reference phantom plots are from stent #6, which was the same stent type as used in patient 3–5. From the patients that received stent #6, none had air in their stent, making a direct comparison to the angulated phantom measurements more challenging. Patient 6, who had air in the stent, received stent #1.

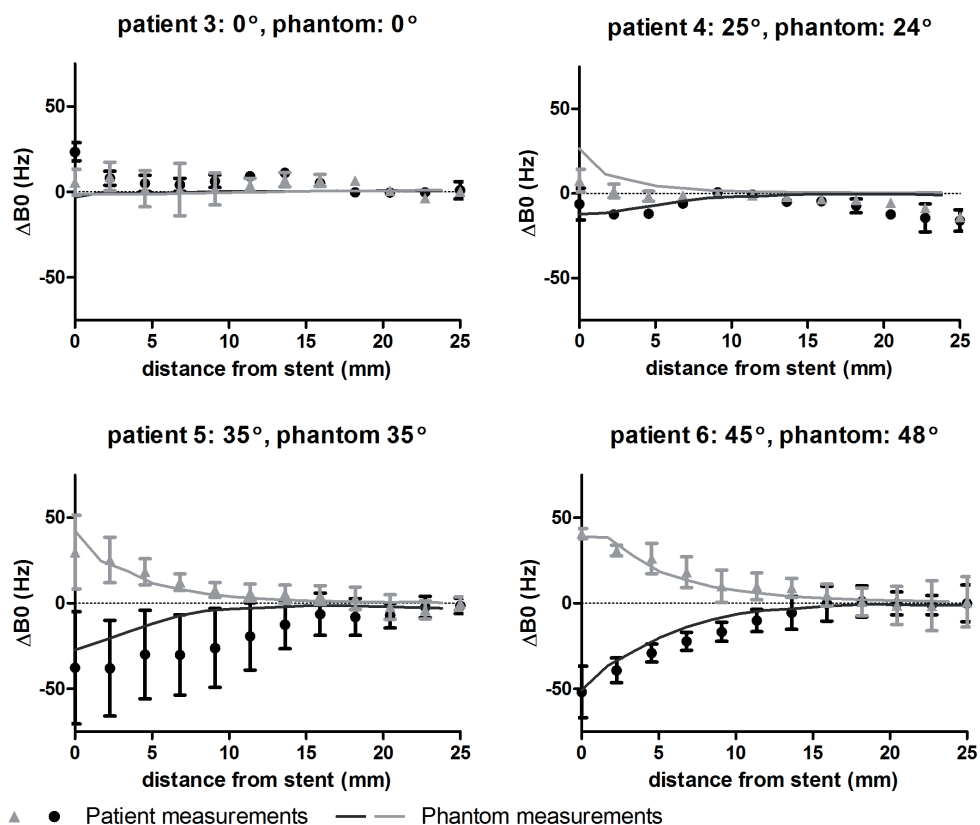


Figure 5.9. Comparison of four *in vivo* measurements (dots and triangles) of stent #1 (patient 6), and #6 (patient 3–5) and the phantom measurements of stent #6 (lines). The lines represent the phantom measurements at a similar angle (data are also shown in Fig. 5.5). ΔB_0 is plotted for two perpendicular lines: in the direction parallel to the projection of B_0 on the slice (circles and dark lines), and the direction perpendicular to the projection of B_0 (triangles and bright lines). The stent in patient 6 contained air. Error bars represent standard deviation over the averaged slices.

Discussion

In this study, we showed that for the eight biliary stents tested, susceptibility artifacts were dominant compared to the gradient-induced artifacts and RF-induced artifacts. We are the first to apply a sequence-independent artifact quantification method [18] to stents and assess B_0 and T_2^* as a function of distance from the stent at 1.5 T as well as 3 T. The produced graphs allow users to calculate the expected artifact for any MRI-sequence. Also, we systematically investigated the dependence of the susceptibility artifacts on the angle between stent and main magnetic field. Finally, similar data acquired *in vivo* showed that for the tested platinum-cored nitinol stents aligned with the main magnetic field, the artifacts remain limited. However, when the stent was at an angle of 30° with the main magnetic field, the effects on B_0 and T_2^* are larger than the effects of typical anatomical susceptibility changes on B_0 and T_2^* . The quantification of the artifacts can be used to determine stent- and MRI sequence-dependent treatment planning margins for target delineation uncertainties in radiotherapy.

Many papers in the literature previously assessed the artifacts caused by stents [21, 22, 25, 38, 49-54], of which two focused on biliary stents [21, 22]. Most papers focused on artifacts inside stents or in the vessel walls [21, 49-52], which is of less importance for radiotherapy. Also, until now, research focused on sequence-dependent measures [21, 22, 25, 38, 49-54] and thus, the results are only valid for the specific sequence settings that were used. In some articles imaging parameters were varied slightly, to allow for a broader applicability (e.g. acquisitions at multiple echo times [54]). However, in none of the publications the underlying mechanisms causing the artifacts were measured, as was done for the susceptibility artifacts in this paper. Instead of measurements, Guo and Jiang performed simulations that used the underlying mechanism of the susceptibility artifacts and RF-induced artifacts for a stent of 40 mm diameter for several sequences [53]. However, these simulations focused on distortions inside the stent and the diameter used was larger than in biliary stents (~ 10 mm). Finally, none of the studies [21, 22, 25, 38, 49-54] mentioned artifact-sensitive sequences, such as DWI.

Our data show that the stainless steel stent exhibits the largest artifacts, whereas the polyethylene stent generates the smallest artifacts. These results are in agreement with previous studies [21, 49]. Even though polyethylene stents are preferable when considering the MR image quality, metal stents are becoming the gold standard due to clinical advantages [55-57].

We focused on a sequence-independent approach to quantify artifacts such that the results can be generalized to any MR sequence. However, there is no sequence- and situation-independent measure for RF-induced and gradient-induced artifacts.

Gradient-induced artifacts depend on the gradient slew rate, the gradient directions, and the location of the stent in the scanner. In this study, we placed the stents as far from the center of the bore as our setup allowed, to maximize eddy currents and hence any gradient-induced artifact. Even at this setting, there was no sign of gradient-induced artifacts. Therefore, in less extreme, clinical, situations, we expect the susceptibility artifacts to be dominating compared to the gradient-induced artifact. These findings are in agreement with previous studies in which gradient-induced artifacts only occurred in setups that were specially designed to induce these artifacts [24, 25]. In addition, the RF-induced artifacts will depend, among other factors, on the quality of the B1-shimming (if multitransmit hardware is available) as well as the location of the stent compared to the transmit channels. The effect of stent position on the RF-induced artifacts was not systematically investigated in this research due to the limited option to reproduce or generalize such results. However, the RF-induced artifacts were strongest inside the stent, which is of less relevance for tumor delineation in radiotherapy.

The stent-related artifacts we assessed are independent of b-value. However, it is known that stent-unrelated eddy currents in the MRI bore can cause b-value dependent deformations of DWI images [58]. These deformations are well understood and can be corrected for [59]. Furthermore, the measured local B_0 inhomogeneity causes a local gradient in B_0 close to the stent, which alters the b-value from its intended value for DWI. The extent of this effect will depend on the specific DWI sequence used.

A limitation of our phantom study is that the stents were fully expanded and filled with agar gel. In patients, stents were deformed and in one of our examples contained air. In fact, MRI data were available for 23 patients that received biliary stents and from these stents, four were filled with air and two contained small air bubbles. Such deformations and air pockets can influence the artifacts *in vivo* and will be strongest for the B_0 inhomogeneity. The agreement of the *in vivo* results with the phantom results in patients 3–5 indicates that deformations had no strong additional effects. From the four patients with air in their stent, none had received stent #6. Therefore, we displayed a patient with stent #1 who had air in the stent. This stent is expected to show smaller artifacts than the stent used as a reference, stent #6 (e.g., Fig. 5.3). However, in our *in vivo* data, we find that the artifacts were of similar magnitude, potentially due to the additional effect of air. From this, we conclude that there may be some additional effect of air pockets in stents, though this was smaller than the effects from the stent itself in our example. As the effect appears to be limited and a minority of patients had air in the stent, we did not further characterize the effect of air. Whether the small number of patients with air in the stent is representative for the

patient group or due to the scans often being obtained after lunch and after drinking two cups of pineapple juice is unclear.

In clinical practice, one could acquire a ΔB_0 -map and T2*-map for each patient to assess patient-specific expected deformations, given the MRI sequence that will be used. Another limitation is the low resolution used for the B1-maps. Therefore, it was often challenging to detect the edge of the stent. However, it is clear that the effect of the stent on B1 was small outside the stent.

In addition to image deformation, signal shifts from susceptibility artifacts may cause hyperintense and hypointense regions when the signal is piled up due to gradients in ΔB_0 . These regions can lead to misinterpretation of the anatomy. Potentially, such a signal pile-up is occurring in both DWI images of patient 1, where a hyperintense region arises anterior to the stent. One can predict the expected deformations and investigate potential signal pile-up using the ΔB_0 -map, Eq. 5.1 (Sec. I of the supplementary material [32]), and a given bandwidth per voxel (BW_{vox}) [53, 60]. This is investigated in more detail in Sec. V of the supplementary material [32].

In this study, no CT images were used as a reference for several reasons. Our *in vivo* MRI images were obtained as part of another study and therefore acquired on average two months before the available CT images. In vivo, stents can migrate over such time periods [30, 61]. Also, variations in stomach and intestinal filling change the shape of the stent between both scans. Therefore, it is not straightforward how to relate the findings on CT to our MRI data and the added value of geometrically accurate CT imaging may be lost in this situation. When MRI is used solely for the purpose of assisting in tumor delineation, it should be obtained directly before or after CT to minimize these effects. For phantom measurements, CT images would need to be registered to the MR acquisitions using, e.g., markers. As these markers would also be susceptible to deformations, there was no strong advantage. Instead, high BW (1602 Hz/voxel) MR images, obtained during the same session, were used. For the highest observed ΔB_0 (100 Hz) in our phantom measurements, excluding the results from the stainless steel stent, this BW results in a maximum shift of 0.06 voxels or 0.1 mm.

From our study, it is clear that for sequences with a higher BW_{vox} (e.g., $BW_{\text{vox}} > 275$ Hz/voxel for 1.5 T and $BW_{\text{vox}} > 500$ Hz/voxel for 3 T), the B_0 -related susceptibility artifacts will be negligible for stents #1–6 and #8 (deformations < 0.2 voxel; see Fig. 5.4 for 1.5 T and Fig. 5.3 for 3 T and Eq. 5.1 from Sec. I of the supplementary material I [32]). When sequences with lower BW_{vox} are desired, such as DWI, images from nonpolyethylene stents must be interpreted with caution. For typical DWI sequences ($BW_{\text{vox}} = 20$ Hz/voxel, voxel size = 2.5 mm) of stents #1–6, signal shifts of up to 1.5 voxels (= 3.75 mm) may occur at a distance of 5 mm from the stent at 3 T.

These stents also showed a decreased $T2^*$ up to 5.1/8.5 mm (1.5/3 T) from the edge of the stent. For the stainless steel stent (#7), the deformations induced at 20 mm distance from the stent are 4 voxels for a typical DWI sequence ($BW_{\text{vox}} = 20 \text{ Hz/voxel}$); with a BW_{vox} of 500 Hz/voxel, deformations of up to 0.5 voxels will still occur at 10 mm from the stent. Therefore, use of stainless steel stents should be avoided if one is interested in reliable MRI data close to the stent.

For target delineation in radiotherapy, uncertainties caused by image deformations can be accounted for by applying local safety margins to ensure irradiation of the target volume. Depending on the proximity of the target volume to the stent, the margin can be adapted according to calculated signal shifts based on Fig. 5.4 (1.5 T) or Fig. 5.3 (3 T) and Eq. 5.1 (Sec. I of the supplementary material [32]). For typical DWI sequences ($BW_{\text{vox}} = 20 \text{ Hz/voxel}$, voxel size = 2.5 mm), and stents included in this research other than the stainless steel stent, signal will shift at most 2.75 voxels (= 6.9 mm), considering a maximum $|\Delta B_0| < 55 \text{ Hz}$ at the stent edge at 1.5 T and 5 voxels (= 12.5 mm) considering a maximum $|\Delta B_0| < 100 \text{ Hz}$ at 3 T. Ultimately, the extra margins can decrease with distance from the stent. When determining these margins, one could include the (sequence specific) direction of expected signal shifts. In the case a ΔB_0 -map is acquired *in vivo*, the shift magnitude could be derived directly from those images instead. As the delineation uncertainty (e.g., due to interobserver variation) on MRI has not been quantified yet, it is unclear how the uncertainty derived in this research compares to the other uncertainties and how this will affect the final margins.

In addition to delineating tumors, MRI may also help in delineating organs at risk. In this case, stents could also cause deformations, especially in the duodenum, which is often in close proximity to the stent. Appropriate safety margins, as derived from our results, can be applied. However, the duodenum often contains air bubbles, which can cause additional B_0 inhomogeneities. Furthermore, organs at risk are generally best visible on high BW sequences, such as the Dixon and T2W-TSE images, which have negligible deformations for the ΔB_0 s induced by the tested stents.

Conclusion

We showed that, of the different MRI artifacts induced by biliary stents, the susceptibility artifacts were dominant over the others. The susceptibility artifacts were quantified independently of the sequence as a function of the distance from the stent. Our findings will contribute to the determination of uncertainties in radiotherapy target delineation, which can be taken into account when establishing safety margins to assure target dose coverage.

References

1. Rasch C, Keus R, Pameijer FA et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 1997; 39(4):841–848.
2. Villeirs GM, van Vaerenbergh K, Vakaet L et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther. Onkol.* 2005; 181(7):424–430.
3. Mitchell DG, Snyder B, Coakley F et al. Early invasive cervical cancer: Tumor delineation by magnetic resonance imaging, computed tomography, and clinical examination, verified by pathologic results, in the ACRIN 6651/GOG 183 intergroup study. *J. Clin. Oncol.* 2006; 24(36):5687–5694.
4. Dempsey JF, Benoit D, Fitzsimmons JR et al. A device for realtime 3D image-guided IMRT. *Int. J. Radiat. Oncol. Biol. Phys.* 2005; 63(2):S202.
5. Raaymakers BW, Lagendijk JJW, Overweg J et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: Proof of concept. *Phys. Med. Biol.* 2009; 54(12):N229–N237.
6. Matsuki M, Inada Y, Nakai G et al. Diffusion-weighted MR imaging of pancreatic carcinoma. *Abdom. Imaging* 2007; 32(4):481–483.
7. Gurney-Champion OJ, Froeling M, Klaassen R et al. Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas. *Invest. Radiol.* 2016; 51(4):211–220.
8. Lemke A, Laun FB, Klaub M et al. Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: Comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. *Invest. Radiol.* 2009; 44(12):769–775.
9. Klaub M, Lemke A, Grünberg K et al. Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. *Invest. Radiol.* 2011; 46(1):57–63.
10. Gaeta M, Benedetto C, Minutoli F et al. Use of diffusion-weighted, intravoxel incoherent motion, and dynamic contrast-enhanced MR imaging in the assessment of response to radiotherapy of lytic bone metastases from breast cancer. *Acad. Radiol.* 2014; 21(10):1286–1293.
11. Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. *J. Magn. Reson. Imaging* 2006; 24(3):478–488.
12. Yamazaki H, Nishiyama K, Tanaka E et al. Dummy run for a phase II multi-institute trial of chemoradiotherapy for unresectable pancreatic cancer: Inter-observer variance in contour delineation. *Anticancer Res.* 2007; 27(4 C):2965–2971.
13. Caravatta L, Macchia G, Mattiucci GC et al. Inter-observer variability of clinical target volume delineation in radiotherapy treatment of pancreatic cancer: A multi-institutional contouring experience. *Radiat. Oncol.* 2014; 9(1):198.
14. Gurney-Champion OJ, Nederveen AJ, Klaassen R et al. Revisiting the potential of alternating repetition time balanced steady-state free precession imaging of the abdomen at 3 T. *Invest. Radiol.* 2016; 51(9):560–568.
15. van der Horst A, Wognum S, Dávila Fajardo R et al. Interfractional position variation of pancreatic tumors quantified using intratumoral fiducial markers and daily cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 87(1):202–208.
16. Dávila Fajardo R, Lekkerkerker SJ, van der Horst A et al. EUS-guided fiducial markers placement with a 22-gauge needle for image-guided radiation therapy in pancreatic cancer. *Gastrointest. Endosc.* 2014; 79(5):851–855.
17. Lens E, van der Horst A, Kroon PS et al. Differences in respiratory-induced pancreatic tumor motion between 4D treatment planning CT and daily cone beam CT, measured using intratumoral fiducials. *Acta Oncol.* 2014; 53(9):1257–1264.

18. Gurney-Champion OJ, Lens E, van der Horst A et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med. Phys.* 2015; 42(5):2638–2647.
19. Porta M, Fabregat X, Malats N et al. Exocrine pancreatic cancer: Symptoms at presentation and their relation to tumour site and stage. *Clin. Transl. Oncol.* 2005; 7(5):189–197.
20. Modolell I, Guarner L, Malagelada JR. Vagaries of clinical presentation of pancreatic and biliary tract cancer. *Ann. Oncol.* 1999; 10(4):82–84.
21. Merkle EM, Boll DT, Weidenbach H et al. Ability of MR cholangiography to reveal stent position and luminal diameter in patients with biliary and in vivo results in 30 patients. *AJR. Am. J. Roentgenol.* 2001; 176(4):913–918.
22. Girard M, Hahn P, Saini S et al. Wallstent metallic biliary endoprosthesis: MR imaging characteristics. *Radiology* 1992; 184(3):874–876.
23. Schenck JF. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Med. Phys.* 1996; 23(6):815–850.
24. Graf H, Steidle G, Martirosian P et al. Metal artifacts caused by gradient switching. *Magn. Reson. Med.* 2005; 54(1):231–234.
25. Shenhav A, Azhari H. Gradient field switching as a source for artifacts in MR imaging of metallic stents. *Magn. Reson. Med.* 2004; 52(6):1465–1468.
26. Vashae S, Goora F, Britton MM et al. Mapping B1-induced eddy current effects near metallic structures in MR images: A comparison of simulation and experiment. *J. Magn. Reson.* 2015; 250:17–24.
27. Graf H, Steidle G, Martirosian P et al. Effects on MRI due to altered rf polarization near conductive implants or instruments. *Med. Phys.* 2006; 33(1):124–127.
28. Lauer UA, Berger A, Schick F. RF artifacts caused by metallic implants or instruments which get more prominent at 3 T : An in vitro study. *Magn. Reson. Imaging* 2005; 23(3):493–499.
29. Camacho CR, Plewes DB, Henkelman RM. Nonsusceptibility artifacts due to metallic objects in MR imaging. *J. Magn. Reson. Imaging* 1995; 5(1):75–88.
30. van der Horst A, Lens E, Wognum S et al. Limited role for biliary stent as surrogate fiducial marker in pancreatic cancer: Stent and intratumoral fiducials compared. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):641–648.
31. Yushkevich PA, Piven J, Hazlett HC et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 2006; 31(3):1116–1128.
32. See supplemental material.
33. Kaye GWC, Laby TH. Tables of Physical & Chemical Constants (16th edition 1995), Longman, 1995.
34. Brudson JA. Plastic Materials, 5th edition. London: Butterworths, 1989.
35. Serway RA, Gordon JR. Principles of Physics, 2nd edition. London: Saunders College Publication, 1998.
36. Lide DR. CRC Handbook of Chemistry Physics, 75th edition. London: CRC, 1994.
37. Antonucci V, Faiella G, Giordano M et al. Electrical resistivity study and characterization during NiTi phase transformations. *Thermochim. Acta* 2007; 462(1–2):64–69.
38. Wang Y, Truong TN, Yen C et al. Quantitative evaluation of susceptibility and shielding effects of nitinol, platinum, cobalt-alloy, and stainless steel stents. *Magn. Reson. Med.* 2003; 49(5):972–976.
39. Gilbert G, Soulez G, Beaudoin G. Improved in-stent lumen visualization using intravascular MRI and a balanced steady-state free-precession sequence. *Acad. Radiol.* 2009; 16(12):1466–1474.
40. Rylander S, Thörnqvist S, Haack S et al. Intensity profile based measurement of prostate gold markers influence on 1.5 and 3T diffusion-weighted MR images. *Acta Oncol.* 2011; 50(6):866–872.
41. Yoshimura K, Kato H, Kuroda M et al. Development of a tissue-equivalent MRI phantom using carrageenan gel. *Magn. Reson. Med.* 2003; 50(5):1011–1017.
42. Kato H, Kuroda M, Yoshimura K et al. Composition of MRI phantom equivalent to human tissues. *Med. Phys.* 2005; 32(10):3199–3208.

43. de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results. *Radiology* 2004; 230(3):652–659.
44. van Heerden LE, Gurney-Champion OJ, van Kesteren Z et al. Quantification of image distortions on the Utrecht interstitial CT/MR brachytherapy applicator at 3T MRI. *Brachytherapy* 2015; 15(1):118–126.
45. Froeling M, Nederveen AJ, Heijtel DFR et al. Diffusion-tensor MRI reveals the complex muscle architecture of the human forearm. *J. Magn. Reson. Imaging* 2012; 36(1):237–248.
46. Nehrke K, Börnert P. DREAM-a novel approach for robust, ultrafast, multislice B₁ mapping. *Magn. Reson. Med.* 2012; 68(5):1517–1526.
47. Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 2012; 9(7):671–675.
48. Wood ML, Henkelman RM. Truncation artifacts in magnetic resonance imaging. *Magn. Reson. Med.* 1985; 2(6):517–526.
49. van der Laan MJ, Bartels LW, Bakker CJG et al. Suitability of 7 aortic stent-graft models for MRI-based surveillance. *J. Endovasc. Ther.* 2004; 11(4):366–371.
50. Spuentrup E, Ruebben A, Mahnken A et al. Artifact-free coronary magnetic resonance angiography and coronary vessel wall imaging in the presence of a new, metallic, coronary magnetic resonance imaging stent. *Circulation* 2005; 111(8):1019–1026.
51. Adams GJ, Baltazar U, Karmonik C et al. Comparison of 15 different stents in superficial femoral arteries by high resolution MRI ex vivo and in vivo. *J. Magn. Reson. Imaging* 2005; 22(1):125–135.
52. Klemm T, Duda S, Machann J et al. MR imaging in the presence of vascular stents: A systematic assessment of artifacts for various stent orientations, sequence types, and field strengths. *J. Magn. Reson. Imaging* 2000; 12(4):606–615.
53. Guo Y, Jiang X. Simulations of the stent artifacts in magnetic resonance imaging. *IEEE Trans. Magn.* 2012; 48(2):659–662.
54. Bartels LW, Smits HF, Bakker CJ, Viergever MA. MR imaging of vascular stents: Effects of susceptibility, flow, and radiofrequency eddy currents. *J. Vasc. Interv. Radiol.* 2001; 12:365–371.
55. van der Gaag NA, Rauws EAJ, van Eijck CHJ et al. Preoperative biliary drainage for cancer of the head of the pancreas. *N. Engl. J. Med.* 2010; 362(2):129–137.
56. Gardner TB, Spangler CC, Byanova KL et al. Cost effectiveness and clinical efficacy of biliary stents in patients undergoing neoadjuvant therapy for pancreatic adenocarcinoma in a randomized controlled trial. *Gastrointest. Endosc.* 2016; 84(3):460–466.
57. Tol JAMG, van Hooft JE, Timmer R et al. Metal or plastic stents for preoperative biliary drainage in resectable pancreatic cancer. *Gut* 2015; gutjnl-2014-308762.
58. Koch M, Norris DG. An assessment of eddy current sensitivity and correction in single-shot diffusion-weighted imaging. *Phys. Med. Biol.* 2000; 45(12):3821–3832.
59. Bodammer N, Kaufmann J, Kanowski M, Tempelmann C. Eddy current correction in diffusion-weighted imaging using pairs of images acquired with opposite diffusion gradient polarity. *Magn. Reson. Med.* 2004; 51(1):188–193.
60. Jonsson JH, Garpebring A, Karlsson MG, Nyholm T. Internal fiducial markers and susceptibility effects in MRI - Simulation and measurement of spatial accuracy. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 82(5):1612–1618.
61. Lee JH, Krishna SG, Singh A et al. Comparison of the utility of covered metal stents versus uncovered metal stents in the management of malignant biliary strictures in 749 patients. *Gastrointest. Endosc.* 2013; 78(2):312–324.

Supplementary Materials

I: Artifacts

Susceptibility artifacts

In susceptibility artifacts, a difference in magnetic susceptibility between tissue and stent locally alters B_0 . Susceptibility artifacts have a two-fold effect on MR image quality. Firstly, the signal from voxels with an altered B_0 is shifted in the reconstructed image. The distance shifted depends on the bandwidth per voxel (BW_{vox}) of the sequence, as well as on the change in B_0 (ΔB_0) induced by the stent. With ΔB_0 expressed in Hz ($\Delta B_0[\text{Hz}] = \gamma[\text{rad/s/T}] \times \Delta B_0[\text{T}]/2\pi$, with γ the gyromagnetic ratio) the signal shift is given by:

$$\Delta x = \frac{\Delta B_0}{BW_{\text{voxel}}} \quad (5.1)$$

Secondly, the local change in B_0 goes alongside with a local spatial slope in B_0 . This slope in B_0 causes intravoxel spin dephasing, which results in a decrease in T_2^* -relaxation time in the vicinity of the stent. Therefore, for T_2^* -weighted sequences, implants may cause hypointense regions in which signal is lost. The degree of signal loss depends on the ratio of the sequence's echo time (TE) and T_2^* value of the tissue:

$$S(TE) = S(0) e^{-\frac{TE}{T_2^*}} \quad (5.2)$$

where $S(TE)$ is the signal at TE and $S(0)$ is the signal for $TE = 0$ ms.

Gradient-induced artifact

Gradient-induced artifacts are caused by short-lasting eddy currents in the stent that occur after sharp gradient switching (typical rise and decay times of 5–200 μs , depending on the stent's conductance [24]). Eddy currents induce local field distortions that cause intravoxel spin dephasing. When eddy currents are not balanced at readout, the intravoxel spin dephasing causes signal loss close to the stent, similar to the effect of T_2^* -decay [24, 25]. The effect of gradient switching on the B_0 around the stent increases when the stent is further away from the center of the scanner bore. Therefore, gradient-induced artifacts increase when the stent is placed further away from the bore center in the direction of the switching gradient. Hence, there is no sequence- or location-independent method to quantify these artifacts.

RF-induced artifact

RF-induced artifacts result from the RF-pulse that causes eddy currents in the stent. These eddy currents induce B_1 heterogeneities that oppose the change in B_1 inside the stent. Therefore, the flip angle is locally altered, resulting in signal variations around the stent [29] and hypointensities inside the stent [38, 52-54] on MR images.

II: Post-processing

Our B_0 -maps and B_1 -maps had gradients in the background that, especially for B_0 -maps, masked the effects of the stents (e.g. Fig. 5.A). To remove these low-frequency background variations in the ΔB_0 -map, we subtracted a low-pass filtered version of the ΔB_0 -map, from the original ΔB_0 -map (Fig. 5.A). The filter used was a 2D Gaussian oriented in a plane perpendicular to the central axis of the stent. The Gaussian had a standard deviation (SD) of 8 voxels in both directions, except for stent #7 for which $SD = 9$ voxel was used as the map showed larger artifacts. The filter size was $3 \times SD$. To ensure this filter would not blur out the artifacts induced by the stent, voxels that were affected by the stent were masked during the creation of the low-pass filtered image (per slice a circular ROI was excluded with radius 4–15 voxels for parallel stent orientation and 6–18 voxel for perpendicular stent orientation; the ROI extended well beyond the tip of the stents). There was no visible contribution of the stent to the background image (e.g. Fig. 5.A).

For the B_1 -map, we used a similar approach (Gaussian kernel size: 4 voxels in both directions, except for stent #7, which had 9 voxels; radius of excluded volume: 3–6 voxels, except for stent #7, which had 10–12 voxels). The $\Delta B_0/T_2^*/B_1$ along several lines were plotted; lines are as indicated in Fig. 5.B.



Figure 5.A. ΔB_0 -map before (left panel) and after (right panel) subtracting the background (middle panel), which was created by low-pass filtering the ΔB_0 -map. The region indicated by green rectangle was excluded when generating the background image.

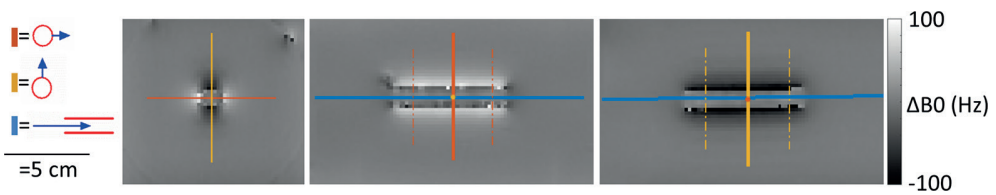


Figure 5.B. Lines for which the data were plotted. Note that for the red and yellow line, the data were only plotted from the edge of the stent onwards. The dashed lines indicate the length over which the data were averaged in the case of the red and yellow lines.

III: Parameter maps

Here we show ΔB_0 -maps (Fig. 5.C), T_2^* -maps (Fig. 5.D) and B_1 -maps (Fig. 5.E) of all stents. The ΔB_0 -map and B_1 -map shown have the background effects filtered out of the image. It is clear that, with exception of stent #7 and #8, the artifacts are of similar magnitude. For all stents, the artifacts focus at the end of the stents in the parallel orientation, whereas the artifacts focus around the edge of the central part of the stent for the perpendicular orientation.

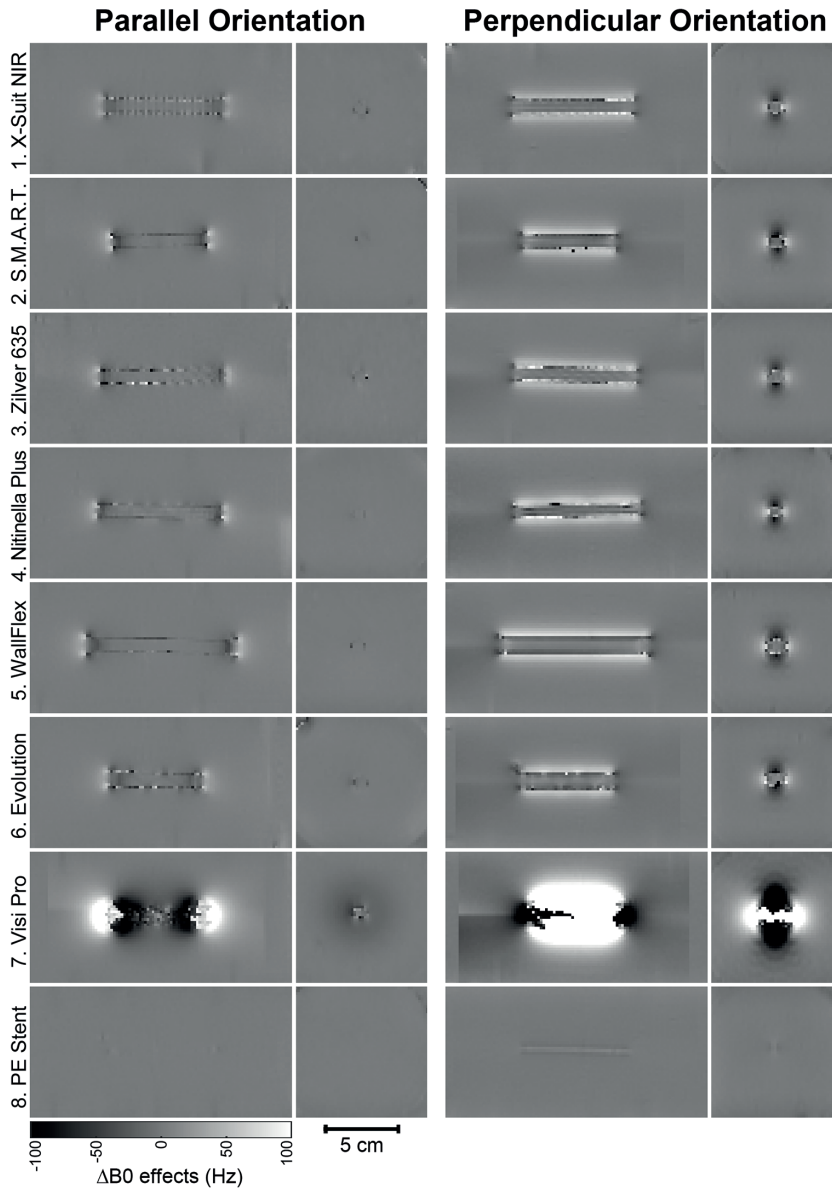


Figure 5.C. ΔB_0 -maps of all stents in parallel (**left**) and perpendicular (**right**) orientation. Note that a different slice orientation is used as in Fig. 5.2, as the represented ΔB_0 -map slices in Fig. 5.2 were aligned with the anatomical slices shown there.

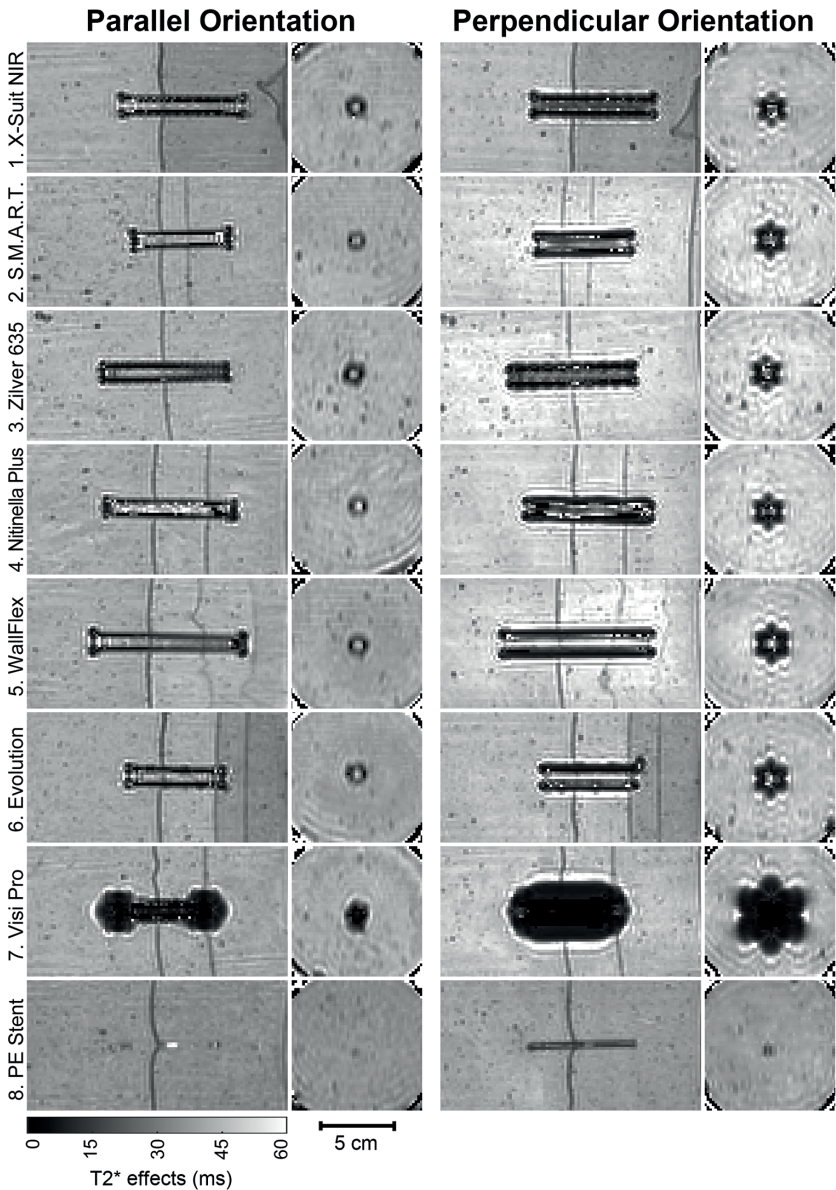


Figure 5.D. T2*-maps of all stents in parallel (**left**) and perpendicular (**right**) orientation. Note that a different slice orientation is used as in Fig. 5.2, as the represented T2*-map slices in Fig. 5.2 were aligned with the anatomical slices shown there.

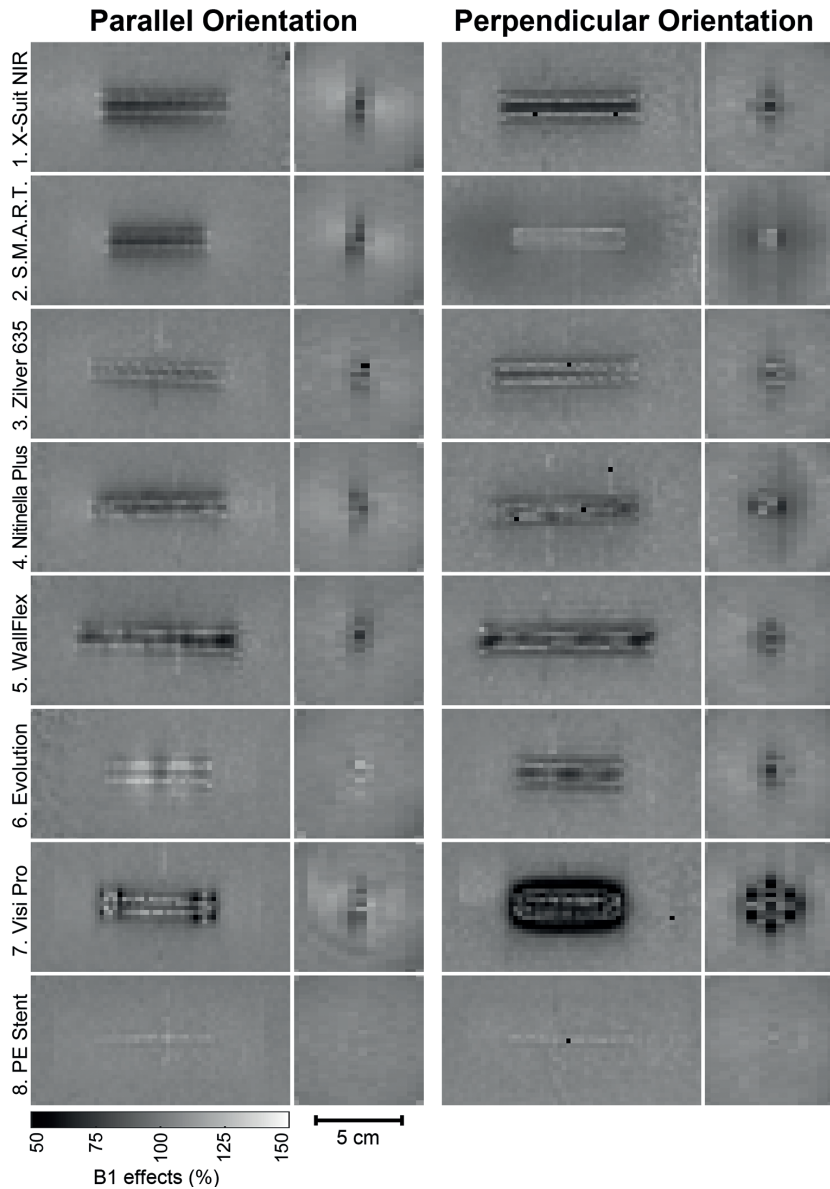


Figure 5.E. B1-maps of all stents in parallel (**left**) and perpendicular (**right**) orientation. Note that a different slice orientation is used as in Fig. 5.2, as the represented B1-map slices in Fig. 5.2 were aligned with the anatomical slices shown there.

IV: Graphs from the phantom results

Here we show the plots of ΔB_0 (Fig. 5.F), T_2^* (Fig. 5.G) and B_1 (Fig. 5.H) as a function of distance for several additional lines in the scans. The bottom row of each figure shows the values plotted for a single line going through the stent (blue line in Fig. 5.B). For treatment planning this is of less interest. Note that for stent #7 the changes in ΔB_0 at some locations changed too rapidly for the dephasing algorithm and, hence, ΔB_0 cannot be trusted inside stent #7. For T_2^* , Gibbs artifacts inside and outside the stent caused the T_2^* fitting algorithm to fail for the affected voxels, yielding artificially high T_2^* values. The B_1 plots show the effect of shielding inside the stent.

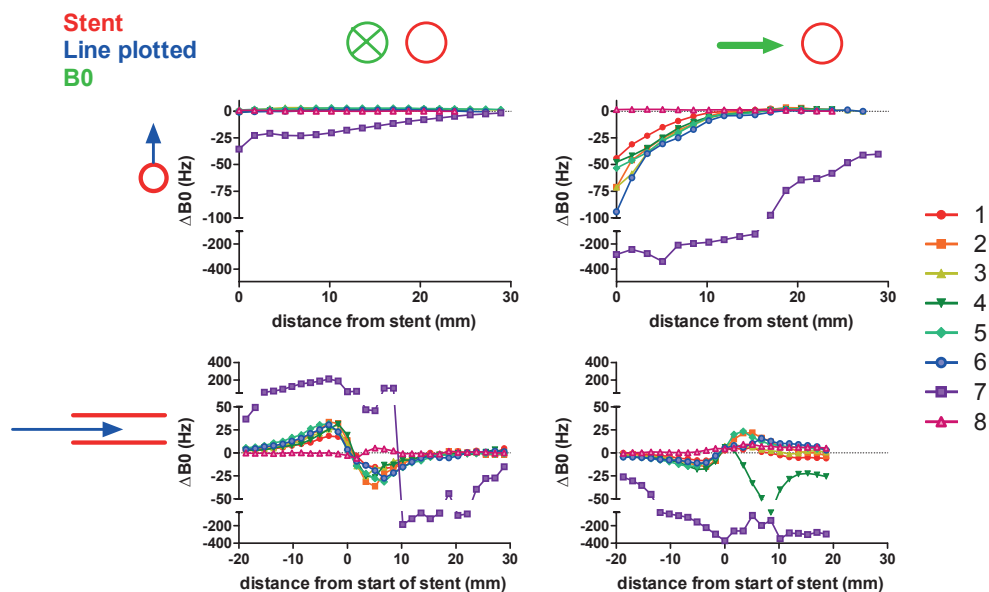


Figure 5.F. Plots of ΔB_0 as a function of distance along several lines (shown left) for the parallel (left column) and perpendicular (right column) orientation. For the line through the stent lumen, all the data have been centered around the leftmost tip of the stent to compare the effects of stents with different lengths.

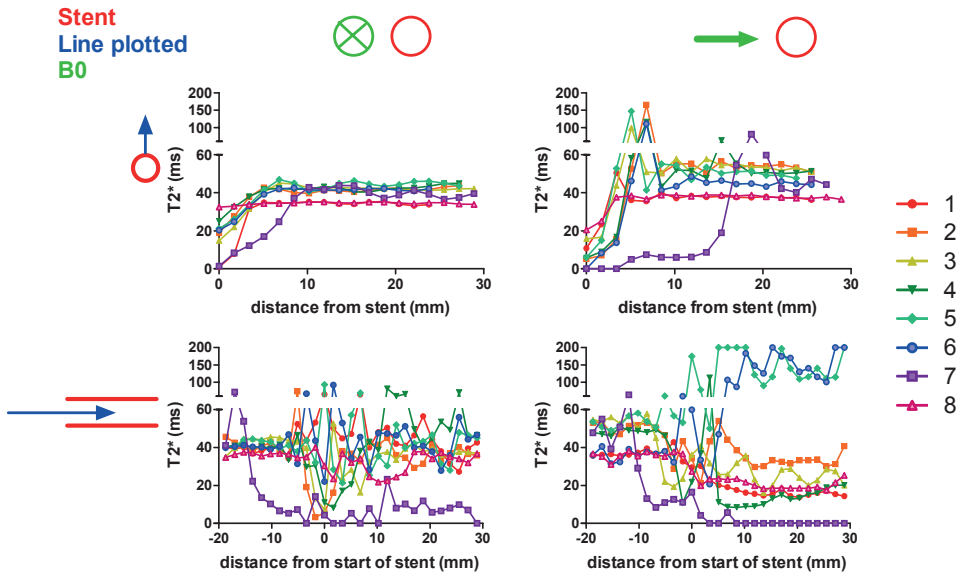


Figure 5.G. Plots of $T2^*$ as a function of distance along several lines (shown left) for the parallel (left column) and perpendicular (right column) orientation. For the line through the stent lumen, all the data have been centered around the leftmost tip of the stent to compare the effects of stents with different lengths.

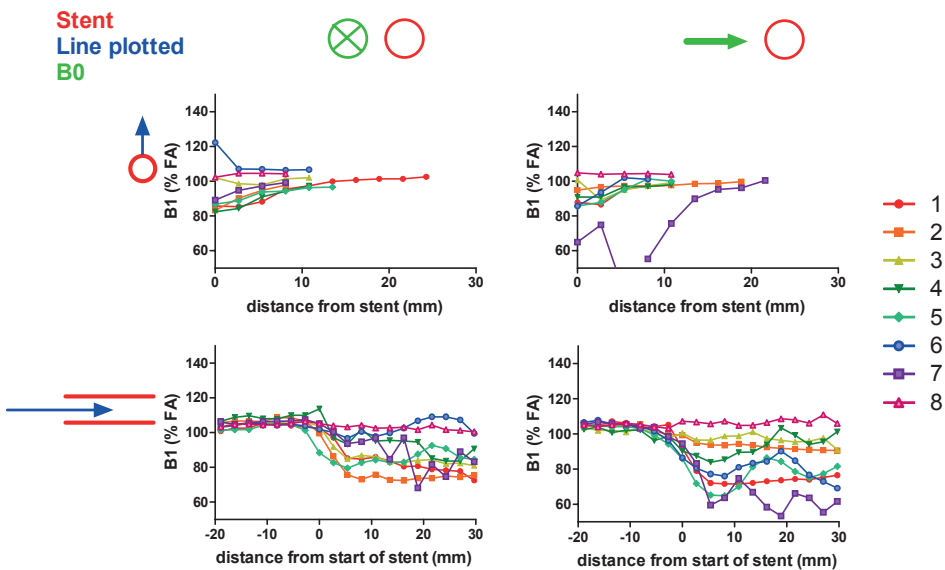


Figure 5.H. Plots of $B1$ as a function of distance along several lines (shown left) for the parallel (left column) and perpendicular (right column) orientation. For the line through the stent lumen, all the data have been centered around the leftmost tip of the stent to compare the effects of stents with different lengths.

V: Simulations

The ΔB_0 -maps give voxel-wise information of the local change in B_0 . When combined with a sequence BW_{vox} , one can predict signal shifts for that sequence. In the discussion, we note that such signal shifts may lead to sequence pile-up that can cause misinterpretation of the images. For example, it is unclear whether the hyperintense regions above the stent on the DWI images of patient 1 (Fig. 5.8, left and middle column) were caused by anatomical structures, or caused by signal shifts. To get an impression of the effect of signal shifts, we simulated signal shifts for a given BW_{vox} . We implemented this simulation by first interpolating the ΔB_0 -map to ten times higher resolution. Then, two equal sized images were created: image A containing ones and image B containing zeros. For each voxel from image A, the signal from its location (x,y) is shifted to location $(x + \Delta x,y)$ in image B, in which Eq. 5.1 determines Δx .¹ the desired BW_{vox} and the ΔB_0 from (x,y) from the high-resolution ΔB_0 -map. Finally, image B was downsampled to the DWI voxel size, in which the voxels from the high-resolution image B that fell within one voxel in the simulated DWI image were summed to form this low-resolution DWI image.

The results of these simulations are shown in Fig. 5.I, in which the signal shifts visible in the simulated images correspond to the signal shifts shown in the actual DWI measurement. Note that in the acquired DW image there is additional signal decay inside the stent due to B_1 -variations and, possibly, T_2 -decay that are not taken along.

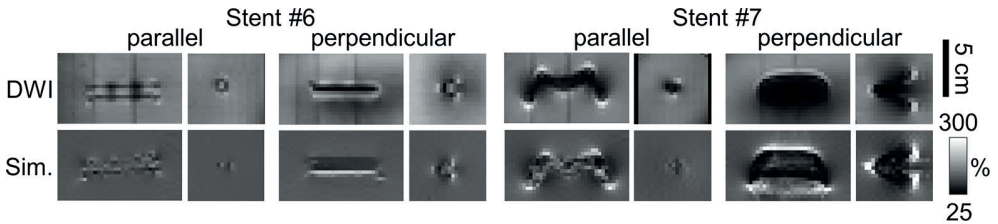


Figure 5.I. Comparison of the artifacts in DWI images ($b = 0 \text{ mm}^2\text{s}$, **top** row) and the simulated artifact from the ΔB_0 -map data using the BW_{vox} from the DWI image (**bottom** row) for stent #6 and stent #7.

For the patients a similar analysis was done on the B_0 -map. Note that in the patient, the homogenous background of ones does not represent patient anatomy. In this case, the simulation can only be used to get an impression of the locations at which signal pile-up may occur. In Fig. 5.J one may appreciate the signal pile-up visible in the simulated image from the central column, which indicates that the hyperintense region in the DWI image (partially) is due to an image artifact. In the left and right column little to no hyperintense and hypointense regions are formed and hence the DWI images can be trusted in these cases.

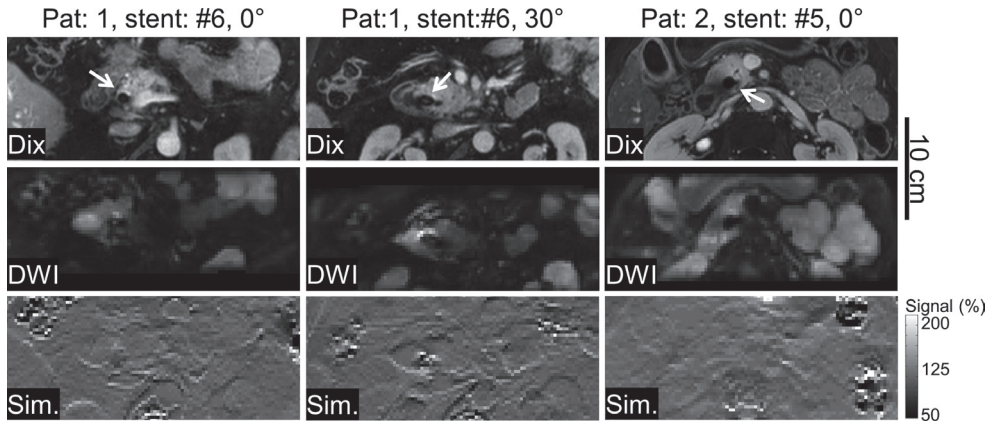


Figure 5.J. Comparison of the DWI images ($b = 0 \text{ mm}^2\text{s}$, **middle row**) and the simulated artifact from the ΔB_0 -map data using the BW_{vox} from the DWI image (**bottom row**). The Dixon images (**top row**) were added as a reference.



CHAPTER 6

ATR-SSFP imaging of the abdomen

Revisiting the potential of alternating repetition time balanced steady-state free precession imaging in the abdomen at 3 T

Oliver J Gurney-Champion

Aart J Nederveen

Remy Klaassen

Marc RW Engelbrecht

Arjan Bel

Hanneke WM van Laarhoven

Jaap Stoker

Sonia I Goncalves

Investigative Radiology 2016;51(9):560–568

Abstract

Objective: The aim was to investigate the value of optimized 3-dimensional alternating repetition time balanced steady-state free precession (ATR-SSFP), as an alternative to conventional segmented balanced steady-state free precession (bSSFP) with fat suppression prepulse (FS-bSSFP), in single breath-hold abdominal magnetic resonance imaging at 3 T.

Methods: Bloch simulations were performed to determine the optimal flip angle ($FA = 1\text{--}90^\circ$) and τ (1–3) with respect to signal-to-noise ratio (SNR) and contrast to noise (CNR) between abdominal organs for ATR-SSFP. These were corroborated by phantom measurements for different $T1/T2$ values (5–47) as well as in a healthy volunteer. In addition, fat suppression efficiency was studied using phantom and volunteer measurements. The effect of resolution on image quality was studied in a healthy volunteer. Using the optimal settings, ATR-SSFP images as well as FS-bSSFP images were obtained in 15 pancreatic cancer patients. For 10 structures of interest, the signal ratio with respect to the pancreas was computed and compared between both sequences. Finally, 10 items on image quality (fat suppression, artifacts, and sharpness) and tissue conspicuity (ducts, vessels and duodenum) were scored by 2 abdominal radiologists for both image sequences.

Results: The results of simulations, phantom measurements, and volunteer measurements showed that, considering scan time, fat suppression, and clinical relevance, the ideal settings for ATR-SSFP were as follows: $\tau = 3$; $TR1 = 3.46$ milliseconds; radiofrequency phase cycling $0^\circ, 180^\circ, 180^\circ, 0^\circ$ and $FA = 13\text{--}16^\circ$ (highest SNR) and $24\text{--}26^\circ$ (highest CNR). The optimized feasible additional settings implemented for patient scans were $FA = 18^\circ$ and resolution = $1.4 \times 1.4 \times 1.4 \text{ mm}^3$. In patients, the signal ratios of both ATR-SSFP and FS-bSSFP were comparable and had a $T2$ -like contrast behavior, although more accentuated in ATR-SSFP. The ATR-SSFP scored significantly higher than FS-bSSFP for 9 of 10 items scored.

Conclusions: For single breath-hold abdominal imaging at 3 T, ATR-SSFP performs best with $\tau = 3$ and a FA between 13° (highest SNR) and 26° (highest CNR). The scoring of both abdominal radiologists indicated that, at $\tau = 3$, $FA = 18^\circ$ and $1.4 \times 1.4 \times 1.4 \text{ mm}^3$ resolution, ATR-SSFP was preferred over conventional FS-bSSFP with similar settings.

Introduction

Balanced steady-state free precession (bSSFP, also known as True FISP, FIESTA, and bFFE) sequences have been progressively applied in abdominal imaging, with and without fat suppression [1-15]. They belong to the category of fast imaging sequences employing steady-state and show a mixed contrast that is weighted on T_1/T_2 . These sequences are very attractive for abdominal imaging not only because they can acquire an entire 3-dimensional (3D) volume in 1 breath-hold, but also because image contrast profits from the broad range of T_1/T_2 values in the abdomen (eg, 6–31) [16]. In addition, bSSFP has the highest signal-to-noise ratio (SNR) among all SSFP sequences because it can theoretically provide up to 50% of the maximum possible signal and has repetition time (TR) typically smaller than 5 milliseconds.

Among the classical clinical applications of bSSFP are cine cardiac imaging [10, 12] and angiography [1, 2, 11, 13]. They both take advantage of the high contrast provided by the different T_1/T_2 of blood and surrounding tissues and benefit from the short acquisition times and high SNR associated with bSSFP. In patients with Crohn disease, 2D/3D fat-saturated and cine bSSFP have been respectively used for improved small bowel wall delineation [3, 4] and to show decreased motility in actively inflamed bowel wall segments [5]. More recently, the use of bSSFP sequences, as an alternative to heavily T2-weighted single-shot (SS) turbo spin-echo (TSE) sequences, was proposed for magnetic resonance cholangiopancreatography (MRCP) [6, 7, 15]. In addition, as fast sequence, bSSFP is useful for real-time guidance of clinical interventions [14]. The contrast in bSSFP images, together with the ability to quickly generate 3D volumes, provides a good overview of the liver and pancreas anatomy as well as a good visualization of the vascular system. It has also been suggested that bSSFP imaging after paramagnetic contrast agent injection has good contrast-to-noise ratio (CNR) for detection of solid lesions in the upper abdomen [8]. Finally, when compared with 2D SS spin echo or 3D TSE, bSSFP is more resistant to the type of flow artifacts that occasionally originate pseudofilling defects in SSTSE images. For that reason, (3D) high-resolution bSSFP can potentially become a valuable additional tool for the visualization of bile and pancreatic ducts and associated pathologies [7]. This is especially relevant for the evaluation of the proximity of a tumor mass to blood vessels [9], the visualization of small biliary stones or mural nodules in pancreatic cystic lesions.

In many of the aforementioned clinical applications, it is desirable to obtain images with fat suppression in order to increase the contrast of the structures of interest, avoid black boundary artifacts due to out-of-phase signal in voxels that contain both fat and water, or to improve the quality of, for example, multiplanar or maximum intensity projection reconstructions. Multiple methods of fat suppression

techniques in bSSFP have been suggested [17-23]; however, the choice is often the use of selective fat saturation prepulses [24]. The introduction of pre-pulses for fat suppression presents several drawbacks. The extra pulses introduce more radiofrequency (RF) power deposition, which may constrain the minimum attainable scan duration, maximum flip angle (FA) or spatial resolution. In addition, the introduction of fat saturation pre-pulses together with the eventual introduction of dummy TRs to minimize signal variations due to the interruption of steady-state will increase scan time. Finally, the periodic interruption of the steady-state introduces variations in the signal that, though minimized by suitable preparation modules [25], modulate the k-space and lead to signal blurring.

A promising bSSFP method that has inherent fat suppression is the alternating TR balanced steady-state free precession (ATR-SSFP) sequence [26]. In ATR-SSFP, the TRs are alternated between TR1 and TR2 while k-space is only acquired during TR1. The alternation between different TRs, together with an adjusted RF pulse phase cycling, modulates the frequency response function (FRF) and creates broad stop-bands positioned at certain dephasing angles (β) that (for a given TR) correspond to well-defined off-resonance spin frequencies. By choosing appropriate parameters, the stop-bands can be used for inherent fat suppression (Fig. 6.1) [26]. Although the introduction of an additional TR also implies an increase in total scan time, when compared with conventional fat-saturated bSSFP, ATR-SSFP does not require the interruption of the steady-state, and, therefore, there is no k-space signal modulation.

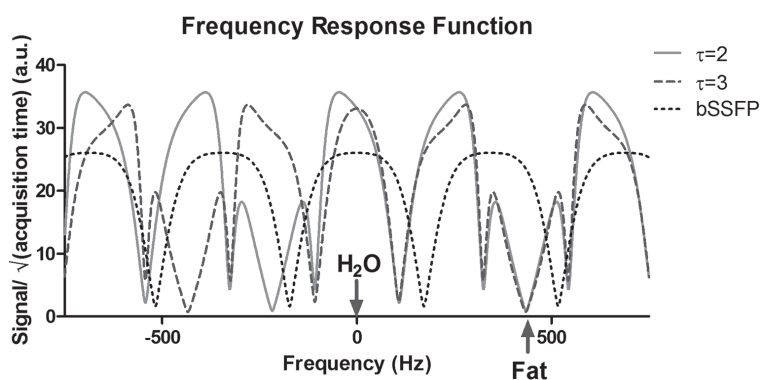


Figure 6.1. Plots of the FRF for bSSFP and 2 of the ATR-SSFP settings that were used in this research. Bloch equations were solved numerically using relaxation values of the pancreas at 3 T: $T_1 = 725$ milliseconds and $T_2 = 43$ milliseconds [16]. Further settings were $B_0 = 3.0$ T and $FA = 20^\circ$ for ATR-SSFP and $FA = 30^\circ$ for bSSFP. Signal intensity has been adjusted for scan time per k-line. Figure 6.1 can be viewed online in color at www.investigativeradiology.com.

Despite these advantages, several limitations associated with ATR-SSFP have prevented its use in routine clinical practice. The ATR-SSFP requires twice as much RF pulses per k-line when compared with conventional bSSFP without fat suppression, which may limit the use of ATR-SSFP for larger FA s, due to specific absorption rate (SAR) constraints. Similarly to bSSFP, ATR-SSFP is sensitive to B_0 heterogeneities that not only result in the appearance of banding artifacts across the image but also decrease the efficiency of fat signal suppression. However, the new-generation magnets have more homogeneous fields and better shimming tools, which makes B_0 heterogeneity less of an issue nowadays.

An additional major factor hindering the routine use of ATR-SSFP is the nontriviality in adjusting imaging protocol parameters when compared with conventional fat-suppressed bSSFP with prepulse (FS-bSSFP). In ATR-SSFP, fat suppression is based on a precise given combination of TR_1 , TR_2 , and RF phase cycling that is field strength dependent. Any change in imaging parameters might render this combination invalid and greatly affect the sequence performance. Furthermore, for abdominal ATR-SSFP, any adjustment in ATR-SSFP imaging parameters should not be detached from the fact that the acquisition time is limited to a breath-hold.

So far, data acquisition strategies for ATR-SSFP have been, to a large extent, chosen heuristically. The FA s ranged from $15\text{--}60^\circ$ and $\tau = TR_1/TR_2$ fluctuated between 1–3 [3, 10, 19, 22, 26–28]. Studies to determine SNR variation in ATR-SSFP focused on relatively small T_1/T_2 ratios, ranging from 1 to 10 [19, 22, 26]. Simulation [22] and phantom results [3] showed that fat suppression efficiency depends on FA and τ . However, no thorough research has been done on the optimal ATR-SSFP settings for abdominal imaging, and no distinct advantage of ATR-SSFP over FS-bSSFP for abdominal imaging has been reported. Therefore, the easier to use FS-bSSFP is currently often used.

The goal of this study is to investigate the value of optimized 3D ATR-SSFP, as an alternative to conventional FS-bSSFP, in abdominal imaging at 3 T. First, we determine the combination of FA and τ that maximizes both signal intensity and contrast using Bloch simulations and experimentally confirm these settings with phantom and volunteer data. Second, we evaluate the value of the optimized ATR-SSFP sequence compared with the conventional FS-bSSFP in patients with pancreatic cancer, and for whom ATR-SSFP could be easily added to their routine imaging protocol.

Methods

Simulations and Phantom measurements

For ATR-SSFP and bSSFP sequences, we ran Bloch simulations using Spin-Bench 1.3 (www.heartvista.com/SpinBench, Los Altos, CA) to predict signal intensity and contrast variation as a function of FA ($1-90^\circ$) for different values of τ (1–3) and $T1/T2$ ratios (3–50) at $B_0 = 3$ T. The simulation settings for ATR-SSFP were: (a) $\tau = 1$, RF phase cycling $0^\circ, 180^\circ, 180^\circ, 0^\circ$ and $TR1 = 3.45$ milliseconds, $\beta = 540^\circ$; (b) $\tau = 2$, RF phase cycling $0^\circ, 240^\circ, 180^\circ, 60^\circ$ and $TR1 = 3.07$ milliseconds, $\beta = 480^\circ$; and (c) $\tau = 3$, RF phase cycling $0^\circ, 180^\circ, 180^\circ, 0^\circ$ and $TR1 = 3.46$ milliseconds, $\beta = 540^\circ$. These settings allowed for the dephasing of fat spins to fall into the stop-band, whereas the nondephasing water spins fall into a pass-band (Fig. 6.1). For conventional alternating phase bSSFP, we set $TR = 3.3$ milliseconds. The echo time (TE) for (ATR-)bSSFP was set to $TE = TR(1)/2$.

We calculated the variation of signal contrast between the pancreas, liver, spleen, and kidney cortex as a function of FA . For this, various $T1/T2$ ratios that are characteristic of different abdominal organs at 3 T were considered: pancreas, $T1/T2 = 16.9$; liver, $T1/T2 = 23.8$; spleen $T1/T2 = 21.8$; and kidney cortex, $T1/T2 = 15.0$ [16]. For simplicity, we assumed similar proton density for all organs. To determine the FA that maximized signal contrast between the pancreas and other abdominal organs, the absolute value of the difference in signal between the pancreas and the different organs was calculated as a function of FA .

To experimentally validate the simulation results and to determine the fat suppression efficiency of ATR-SSFP as a function of FA and τ , we carried out phantom measurements. The phantom, similar to the one used in Coolen et al. [29], consisted of 15 tubes ($\varnothing = 30$ mm, length = 115 mm) that hung in a rack that was inside a container filled with water. Thirteen of the tubes contained agarose gels that consisted of a mixture of 0–45 g of agarose, 10–500 μmol of GdCl_3 , 30 g of carrageen, 1.48 g of NaCl and 3 g of NaN_3 per liter of gel [30]. These gels had $T1/T2$ ratios ranging from 5–47. The central tube was filled with peanut oil, which was taken as a surrogate for human fat.

All measurements were obtained on an Ingenia 3 T scanner (Philips Healthcare, Best, the Netherlands), using a 16-channel phased-array anterior coil and a 10-channel phased-array posterior coil. We measured the $T1$ values of the tubes using the Look-Locker [31] method and the $T2$ values with a Carr-Purcell-Meiboom-Gill acquisition [32, 33].

To partially overcome SAR constraints, a sinc pulse with 1 zero crossing at each side was used for magnetization excitation in the phantom and volunteer

measurements. In this situation, the largest FA in the ATR-SSFP sequence was still limited to $FA = 51^\circ$, 35° , and 21° for ATR-SSFP with $\tau = 1$, $\tau = 2$, and $\tau = 3$ respectively. The 3D scans were acquired with an oversampling factor of 1.6 in the slice direction to account for spurious excitations from outside the field of view (FOV) due to the imperfect pulse slab profile.

For the phantom measurements with ATR-SSFP and bSSFP, parameters τ , $TR(1)$, TE , and RF phase cycling, values were chosen equal to those used in the simulations. Other acquisition parameters for ATR-SSFP ($\tau = 1, 2$, and 3) and bSSFP were as follows: FOV = $300 \times 200 \times 250 \text{ mm}^3$, resolution = $2.5 \times 2.5 \times 2.5 \text{ mm}^3$, sagittal slices, and parallel imaging (SENSE) factor of 2 in the anterior–posterior (AP) direction. For each experimental setting, images were acquired at different excitation FA s (1° , maximum possible FA ; steps of 2°) within a single dynamic acquisition loop to avoid different scaling factors between images. A B1-map was acquired to correct the spatial variation of FA values due to imperfections in B1-shimming.

We analyzed the phantom data in Wolfram Mathematica 9.0 (Wolfram Research Inc, Oxfordshire, United Kingdom). Per tube, we calculated the mean signal within a region of interest (ROI) as a function of FA . The ROIs were cylinders, centered on the longitudinal axis of the tube, which spanned 5 slices and had a volume of 2.5 cm^3 .

Fat suppression efficiency in ATR-SSFP was determined by calculating the ratio between the average signal from all ROIs except the one containing peanut oil and the signal from the ROI in the tube containing peanut oil. The fat suppression efficiency was compared with that of the vendor supplied 3-point Dixon reconstruction that is currently used for pancreatic research at our hospital (Table 6.1).

Table 6.1. In vivo sequence settings.

	ATR-SSFP ($\tau = 3$)	FS-bSSFP	T2W TSE	Dixon
FOV, mm ²	400 × 272	400 × 272	400 × 369	400 × 350
Resolution, mm ²	1.4 × 1.4	1.4 × 1.4	1.3 × 1.6	1.7 × 1.7
Slice thickness, mm	1.4	1.4	2	1.7
slices	64	64	45	53
$TR/TE/\Delta TE$, ms	3.4/1.7/—	2.9/1.5/—	779/80/—	4.4/1.15/0.9
FA /refocussing FA , °	18/NA	30/NA	90/120	10/NA
SENSE	2.4/1.5 (AP/FH)	2.4/1.5 (AP/FH)	2 (AP)	2/1.5 (LR/AP)
Fat-saturated	FRF dependent	SPIR	SPAIR	3-point Dixon

Abbreviations: ATR-SSFP = alternating repetition time balanced steady-state free precession; FS-bSSFP = fat-suppressed bSSFP with prepulse; T2W TSE = T2-weighted turbo spin-echo; FOV = field of view; TR = repetition time; TE = echo time; ΔTE = the increase in echo time in the Dixon sequence; FA = flip angle; NA = not applicable; AP = anterior–posterior; FH = foot–head; LR = left–right; FRF = frequency response function; SPIR = spectral presaturation with inversion recovery prepulse; SPAIR = spectral attenuation inversion recovery prepulse.

In vivo optimization

The optimal ATR-SSFP settings for signal contrast and maximum attainable spatial resolution retaining diagnostic value were tested in 2 healthy volunteers (1 male, 1 female; 22 and 24 years old, respectively) who gave written informed consent.

Three-dimensional volume images were acquired in healthy volunteer 1 using ATR-SSFP and bSSFP at various FAs , and considering different τ values in the case of ATR-SSFP. The settings for ATR-SSFP and bSSFP were similar to those used in simulations. Further acquisition parameters were as follows: FAs in the range 5–55° for $\tau = 1$ and conventional bSSFP, 5–35° for $\tau = 2$, and 5–20° for $\tau = 3$, FAs increased in steps of 5°; FOV = $360 \times 270 \times 90 \text{ mm}^3$, resolution = $2.5 \times 2.5 \times 2.5 \text{ mm}^3$, transverse slices, partial Fourier factor of 0.8, and no parallel imaging. The scan time was 11.4 seconds for conventional bSSFP, 15.2 seconds for ATR-SSFP with $\tau = 2, 3$, and 22.7 seconds for ATR-SSFP with $\tau = 1$. For each τ value, we acquired noise images obtained by running the sequence without RF pulses, as well as data corresponding with the different FAs within 1 dynamic loop. Each image from the loop was acquired in a separate breath-hold.

For data analysis, we selected 1 slice at the center of the imaging volume where we chose ROIs that contained homogeneous signal corresponding either with pancreas, liver, spleen, kidney, or fat. The ROIs were 1.0 cm^2 , except for the ROI in fat, which was smaller (0.5 cm^2). The SNR was calculated as the ratio between the mean signal intensity of the anatomical ROI and noise estimation. The noise estimation was calculated as the standard deviation of a large (150 cm^2) ROI in the noise image, divided by 0.66 [34].

To determine the highest feasible resolution of a 3D isotropic ATR-SSFP volume acquisition within 1 breath-hold (20 seconds), we acquired ATR-SSFP scans at 4 different resolutions in healthy volunteer 2. To keep acquisition time within 1 breath-hold, we increased the SENSE factor at higher resolutions. All 4 (I–IV) scans were obtained using the ATR-SSFP $\tau = 3$ protocol. Additional scan acquisition parameters were as follows: FOV = $400 \times 272 \times 90 \text{ mm}^3$, transverse slice orientation, resolution = $2 \times 2 \times 2$ (I)/ $1.7 \times 1.7 \times 1.7$ (II)/ $1.4 \times 1.4 \times 1.4$ (III)/ $1.2 \times 1.2 \times 1.2$ (IV) mm^3 , $FA = 20^\circ$ and a SENSE factor of 1.8(I)/2.1(II)/2.4(III)/2.5(IV) in the AP direction, and 1(I)/1.2(II)/1.5(III)/2(IV) in the foot–head direction. An abdominal radiologist (J. Stoker), with 20 years of experience, qualitatively analyzed the images. In this analysis, the radiologist focused on the potential diagnostic value, which was assessed by evaluating the following aspects: SNR, the presence of artifacts, and the quality of anatomical details.

Patient measurements

For patient data acquisition, using ATR-SSFP with $\tau = 3$, we selected the RF pulse with 2 zero crossings which allowed a maximum FA of 18° . This FA was still close to the optimal FA value that was found from simulations and confirmed *in vivo*, but the RF pulse provided a better slab profile.

To compare the performance of ATR-SSFP with that of FS-bSSFP in a patient group, we acquired both scans in 15 patients with pancreatic cancer, undergoing different treatments in our hospital and for which both sequences could be added to the corresponding imaging protocol. Patients, all of whom gave written informed consent, underwent magnetic resonance imaging (MRI) scanning as part of ongoing studies (NCT01989000, NCT01995240, NCT02358161), which were all approved by the ethics committee of our institute. A subgroup of 9 patients received hyoscine bromide (Buscopan, 20 mg intravenous) 1–3 minutes before both the ATR-SSFP and FS-bSSFP scans.

Both ATR-SSFP and bSSFP are sensitive to B_0 heterogeneities. Therefore, all patients drank 2 cups of pineapple juice before the MR session that filled the stomach and duodenum and minimized air-tissue boundaries in the FOV. In addition, shimming was focused on the pancreas. We made sure to use implants with small impact on the B_0 homogeneity for the typical metal implants seen in our patient group (biliary stents and fiducial gold markers that are used as guidance for radiotherapy) [35].

In addition to 3D ATR-SSFP and FS-bSSFP images, we acquired 3D Dixon and 2D SS T_2 -weighted (T_2W) TSE images as a reference (Table 6.1). The settings for the FS-bSSFP were similar to the parameters used typically in abdominal imaging at our institute; however, the major parameters (eg, resolution, FOV, SENSE acceleration factor) were kept comparable to ATR-SSFP for more direct comparison (Table 5.1). The segmented FS-bSSFP acquired 21 k-space lines per segment, with an $\alpha/2$ pulse combined with 5 dummy TR s as a startup for each shot. Due to SAR and timing constraints, we chose a FA (30°) that was lower than the simulated optimal values considering CNR (42 – 47°) but similar to the ideal FA considering SNR (23° – 29°). The FA used for ATR-SSFP was also closer to the FA of simulated maximum SNR than the FA of simulated maximum CNR. Shimming was identical for ATR-SSFP and FS-bSSFP.

A researcher (R. Klaassen) placed ROIs in the pancreas, liver, spleen, kidneys, duodenum, pancreatic duct, bile duct, cyst, aorta, inferior vena cava, and fat, provided the structures were present and visible. The signal relative to the pancreas was calculated as

$$S_{rel} = \frac{1}{N} \sum_{n=1}^N \frac{S_{n,tissue}}{S_{n,pancreas}} \quad (6.1)$$

where $S_{n,\text{tissue}}$ and $S_{n,\text{pancreas}}$ were the mean signal from the ROI delineating the tissue of interest and pancreas of patient n , respectively, and N was the total number of patients.

In addition, the image quality of ATR-SSFP and FS-bSSFP images was scored by 2 abdominal radiologists (M.R. Engelbrecht and I. Somers, 8 and 1.3 years of experience, respectively), separately. The radiologists received Dixon and T2W TSE images as reference and were blinded to the sequence they were scoring. Images were classified on a scale from 0 (worst) to 4 (best), in which 0 meant the image had no clinical value, 2 was clinically acceptable, and 4 corresponded with an extremely good image quality. Per patient, both radiologists scored for: fat suppression (excluding subcutaneous fat) as well as image sharpness and artifacts (excluding failure of fat signal suppression and artifacts in the bowel region, as patients were not prepared for bowel MRI). The radiologists were also asked to score the visibility and conspicuity of structures of clinical interest, in which scores were as follows: 0, undetectable; 1, poorly visible; 2, visible, but not delineable; 3, delineable, but without details; 4, details are delineable. The selected structures of interest were as follows: main pancreatic duct, pancreatic duct side branches, celiac trunk, superior mesenteric artery, superior mesenteric vein, duodenum, and gastroduodenal artery.

The agreement between both radiologists was determined by calculating the interclass correlation coefficient (ICC), where an ICC less than 0.4 was considered poor, 0.4–0.6 fair, 0.6–0.75 good, and more than 0.75 excellent. After averaging the results over both radiologists, we checked for significant ($p < 0.05$) differences between the scoring for ATR-SSFP and FS-bSSFP using a Wilcoxon signed-rank test for each item.

Results

Simulations and Phantom measurements

Results from the simulations and phantom measurements are shown in Fig. 6.2 a-h. The FA that allowed highest signal from the liver, kidneys, pancreas, and spleen for ATR-SSFP was in the range 16° – 20° when $\tau = 1$, 15° – 18° when $\tau = 2$, and 13° – 16° when $\tau = 3$ (results not plotted). This was lower than for conventional bSSFP, where FA ranged from 23° – 29° . The ratio between maximal signal magnitudes and the square root of the scan time per k-line was similar ($< 4\%$ difference) for the 3 ATR-SSFP scans and the bSSFP scan. Simulations also showed that the FA that allowed for the highest contrast between the pancreas and its neighboring organs ranged between

30° and 32° when $\tau = 1$, 28°–30° when $\tau = 2$ and 24°–26° when $\tau = 3$ for ATR-SSFP (Fig. 6.2 d-f), and 42° and 47° for conventional bSSFP (Fig. 6.2 h). For $\tau = 3$, the contrast for the liver, spleen, and kidney were optimal at $FA = 22^\circ\text{--}24^\circ$, $FA = 22^\circ\text{--}25^\circ$ and $FA = 24^\circ\text{--}26^\circ$, respectively (results not shown). Phantom measurements agreed with the simulations (Figs. 6.2 a-c and g) for all sequence configurations.

For the phantom measurements, fat suppression efficiency of ATR-SSFP was better than that of Dixon reconstructed images for a broad range of FAs (Fig. 6.2 i).

***In vivo* optimization**

Data from volunteer 1 showed that the SNR was optimal with $FA = 10\text{--}20^\circ$ for $\tau = 1$, $FA = 10\text{--}25^\circ$ for $\tau = 2$ and $FA = 10\text{--}15^\circ$ for $\tau = 3$ in ATR-SSFP and with $FA = 15\text{--}30^\circ$ for bSSFP (Fig. 6.3). These ranges approximately correspond with the range of FA values in which the simulated signal magnitudes were maximal for the organs investigated. Residual variations are probably related to differences between T_1 and T_2 values *in vivo* compared with the values used in the simulations and phantom experiments. Fat suppression was more efficient for $\tau = 2$ and $\tau = 3$ than for $\tau = 1$. Signal-to-noise ratio per square root of acquisition time was similar for all sequences, which agrees with the simulation results.

In healthy volunteer 2, we tested the effect of increasing spatial resolution on the quality of the images (Fig. 6.4). The abdominal radiologist confirmed that the degree of anatomical details and the potential diagnostic value improved at higher resolutions down to 1.4 mm isotropic. The SNR from the data obtained with 1.2 mm isotropic resolution was still deemed suitable for diagnostic purposes. However, at 1.2 mm isotropic resolution, artifacts, such as ghosting of the aorta and SENSE artifacts, became too prominent and dominated the potential gain of the higher resolution. Considering these issues, we selected 1.4 mm isotropic resolution as the optimal setting for ATR-SSFP imaging of the pancreas.

Considering the results of the phantom and *in vivo* measurements, and the FA limit for a pulse with 2 zero crossings, we selected $\tau = 3$, $FA = 18^\circ$ and resolution of 1.4 mm isotropic as optimized settings to acquire ATR-SSFP in patients. These yielded the best compromise between scan time duration, fat suppression efficiency, SNR, and CNR while using an RF excitation pulse that provided a good slab profile.

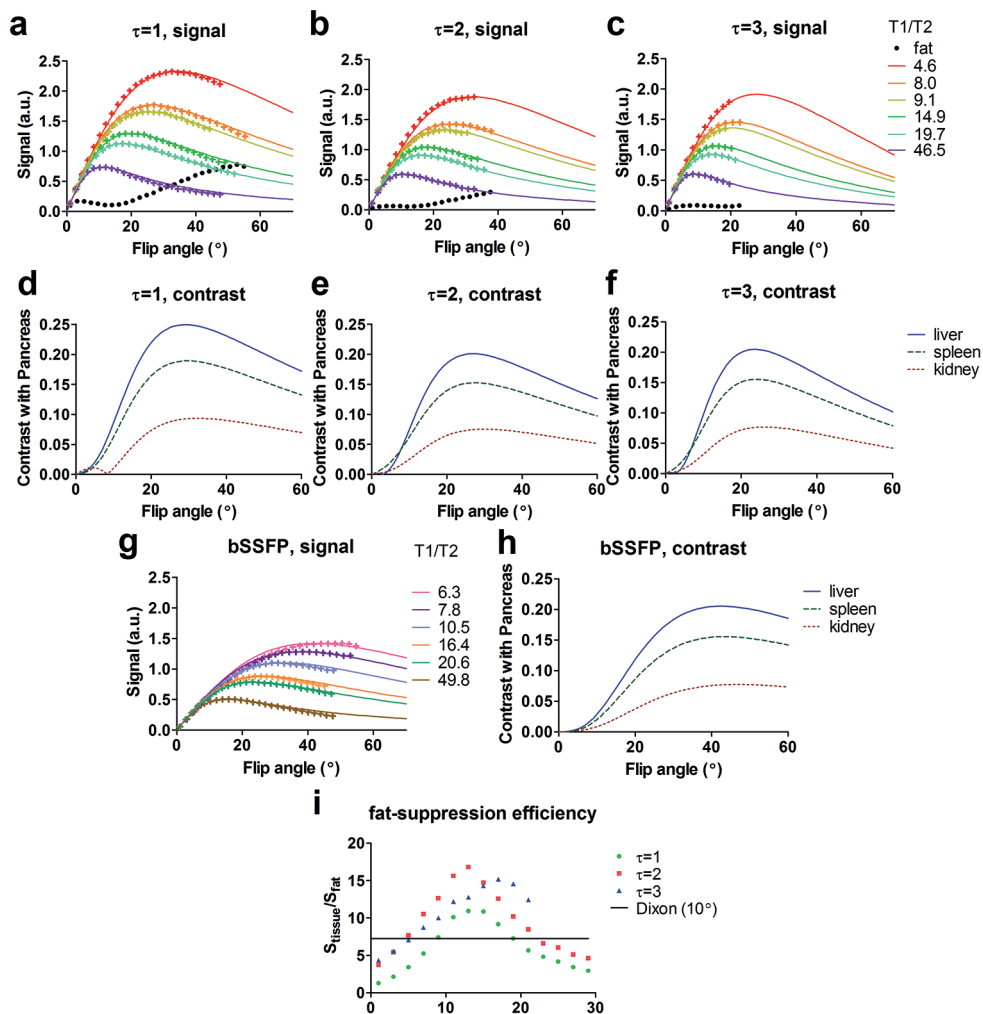


Figure 6.2. Signal variation as a function of FA for several T1/T2 ratios for ATR-SSFP (a-c) and bSSFP (g) for simulations (lines) as well as phantom measurements (dots). Simulated contrast variation between different abdominal organs and the pancreas, as a function of FA for ATR-SSFP (d-f) and for conventional bSSFP (h). Fat suppression efficiency of ATR-SSFP as function of FA as obtained from phantom measurements (i). The maximum allowed FA decreased with increasing τ due to SAR constraints. The change in the T1/T2-ratios from the legend of a to c and g is explained by the aging of the phantom.

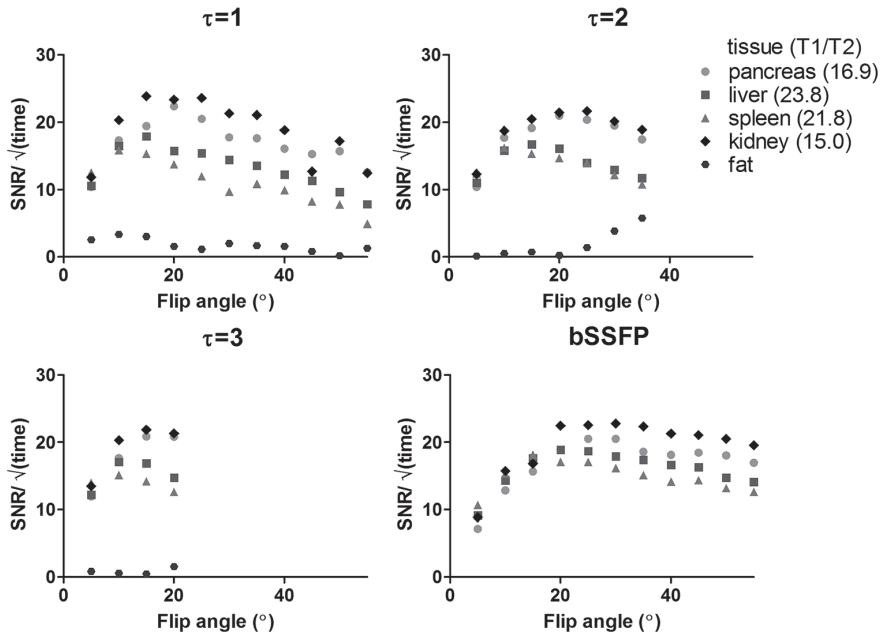


Figure 6.3. *In vivo* SNR (normalized for scan time) as a function of FA for ATR-SSFP with different τ and conventional bSSFP. Note that fat is not plotted in the conventional bSSFP graph as it reached an SNR/√(time) of 180 and was above 30 for all FA > 5°. Figure 6.3 can be viewed online in color at www.investigativeradiology.com.

Patient measurements

The ATR-SSFP and FS-bSSFP images were acquired in all 15 patients (Fig. 6.5). The mean ROI size to determine signal intensity was 66 voxels (range, 22–128 voxels). The contrast in ATR-SSFP and FS-SSFP were similar and mimicked the contrast of T2W TSE scan for most tissues (Figs. 6.5 and 6.6 a). Flowing blood, however, was hyperintense on ATR-SSFP and hypointense on T2W TSE. Blood is expected to be hyperintense on T2-weighted scans due to its long T_2 [36]. However, in T2W TSE, the signal from blood is lost as blood flows out of the slice during the TSE acquisition.

The image scoring agreement between both abdominal radiologists was good (mean ICC = 0.63), and ATR-SSFP scored significantly higher on all items, except for fat suppression, where the scoring difference between ATR-SSFP and FS-bSSFP was not significant, and ICC was lowest (Fig. 6.6 b).

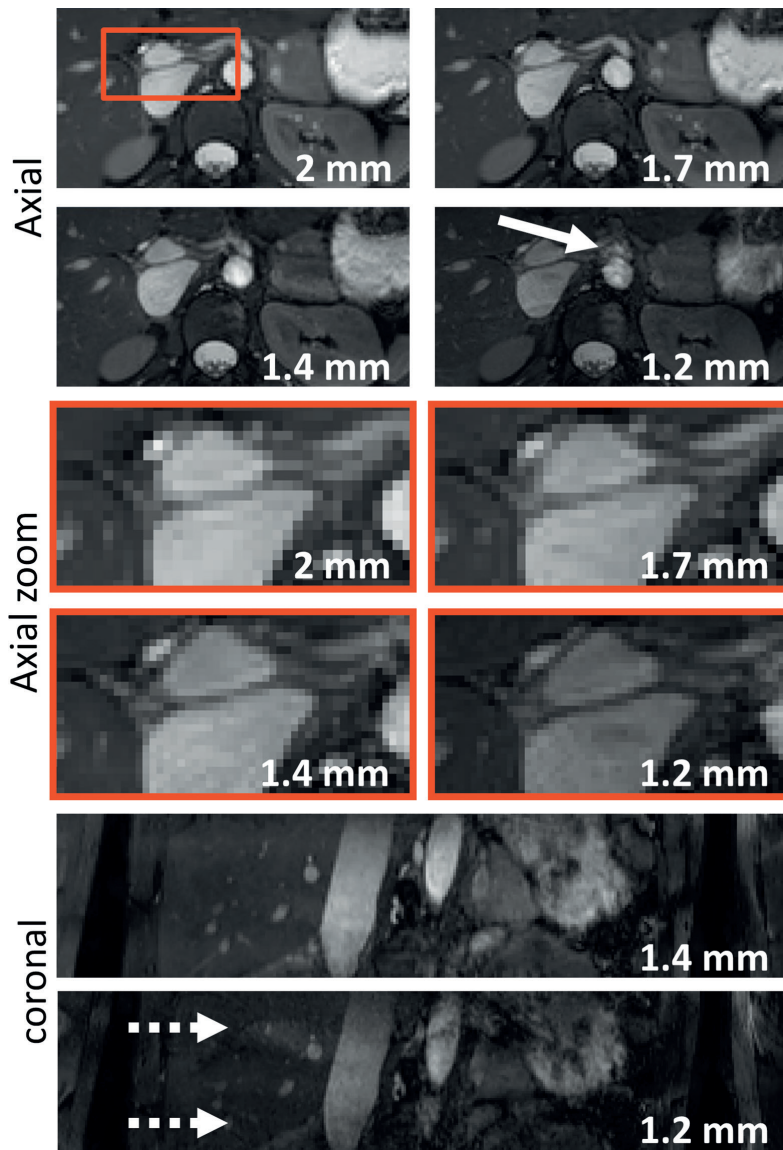


Figure 6.4. *In vivo* ATR-SSFP measurements at 2, 1.7, 1.4, and 1.2 mm isotropic resolution, with $\tau = 3$ and $FA = 20^\circ$, obtained in healthy volunteer 2. At 1.2 mm, there was additional noise due to high SENSE factors (dotted arrows) and pulsation artifacts of the aorta (solid arrow). Figure 6.4 can be viewed online in color at www.investigativeradiology.com.

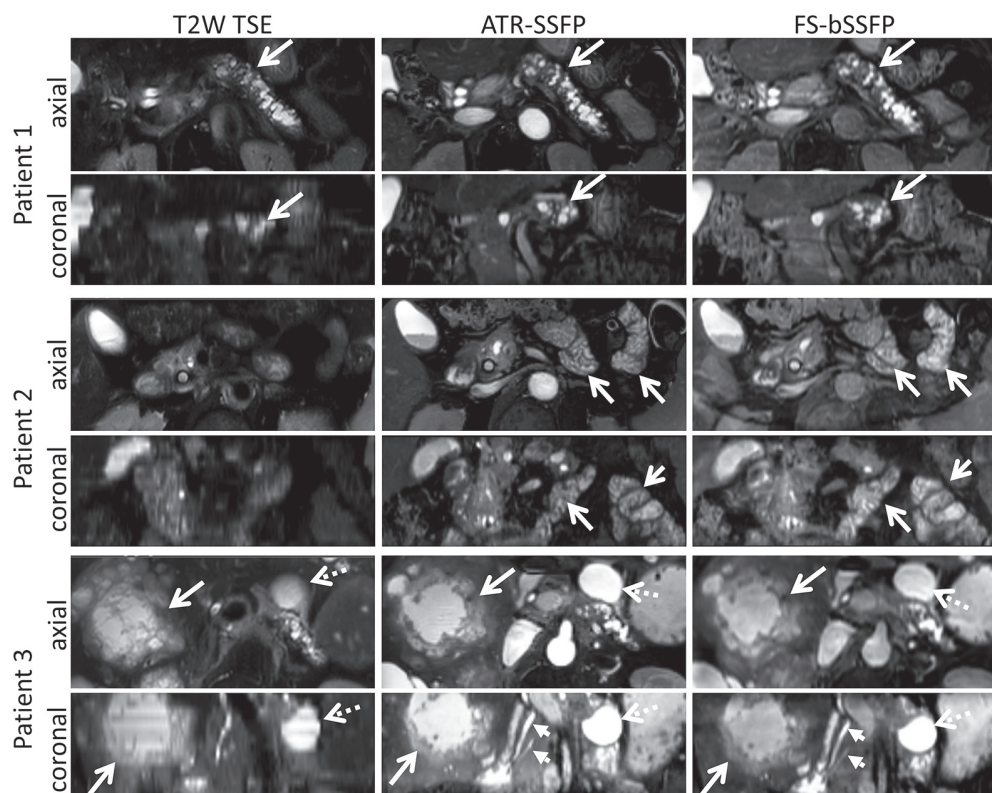


Figure 6.5. Typical examples of T2W TSE (left column), ATR-SSFP (mid column), and FS-bSSFP (right column) images in 3 patients. For each patient, we show the axial and reconstructed coronal views. Arrows indicate structures of interest: dilated pancreatic duct (patient 1), intestines (patient 2), necrotic liver metastasis (patient 3, solid arrow), cyst (dotted arrow), and 2 pancreatic ducts that were not as clearly depicted on the T2W TSE (arrowheads).

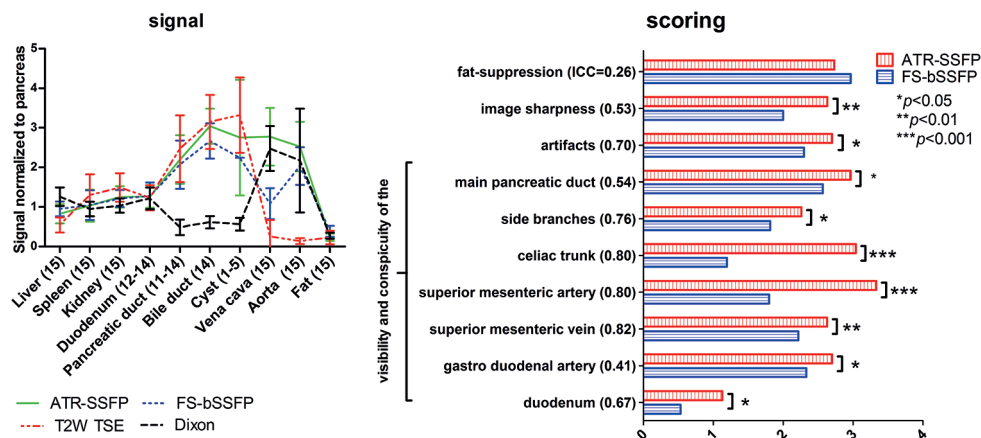


Figure 6.6. Patient results. **Left,** Plot of the relative signal in various tissues with respect to the signal from the pancreas, for various sequences. The numbers on the horizontal axis indicate in how many patients the tissue was delineated (range over sequences). The bars indicate the standard deviation over patients. **Right,** The mean scores of ATR-SSFP and FS-bSSFP, averaged over patients and observers, and corresponding standard deviations. For each item, the ICC is written within brackets.

Discussion

This is the first study demonstrating the value of optimized 3D ATR-SSFP, as an alternative to conventional FS-bSSFP, in abdominal imaging at 3 T. At optimal settings, ATR-SSFP significantly outperforms FS-bSSFP on 9 of 10 scored items. We showed that, for our scanner, the feasible optimal settings for the parameters of single breath-hold ATR-SSFP of the abdomen at 3 T, with respect to fat suppression effectiveness, SNR, CNR between abdominal tissues, and potential diagnostic value, are $FA = 18^\circ$, $\tau = 3$, $TR1 = 3.46$ milliseconds, RF phase cycling $0^\circ, 180^\circ, 180^\circ, 0^\circ$ and an isotropic resolution of $1.4 \times 1.4 \times 1.4 \text{ mm}^3$.

At optimized settings, ATR-SSFP has a similar contrast to the clinically used FS-bSSFP. This T2-like contrast resembles the high contrast between abdominal tissue types that is characteristic of T2W TSE images. In addition, the fat suppression, which is inherent to ATR-SSFP, is of similar quality as conventional FS-bSSFP.

However, ATR-SSFP has advantages over FS-bSSFP. The higher signal in vessels, higher image sharpness, and fewer artifacts of ATR-SSFP compared with FS-bSSFP contributed to a higher scoring in vessel conspicuity (Fig. 6.6) and render ATR-SSFP a good candidate for angiography. In addition, the increase in vessel conspicuity is valuable when considering the assessment of the proximity of a tumor mass to blood vessels. The increased image sharpness, as well as the better visibility and conspicuity

of the main pancreatic duct and its side branches, indicate the potential additional value of ATR-SSFP for the visualization of small biliary stones and mural nodules in pancreatic cystic lesions or obstruction of the ducts in pancreatic tumor patients. Likewise, the clear visibility of the ducts renders ATR-SSFP a good candidate for MRCP, which is currently done with FS-bSSFP [6, 7]. ATR-SSFP enables to obtain MRCP-like images in a single breath-hold and allows for sharper vessel delineations than on FS-bSSFP. Though it was not scored, the sequence gave high bowel conspicuity (Fig. 6.5, patient 2) in the subgroup of patients that got injected with hyoscine bromide before the ATR-SSFP and FS-bSSFP scans.

The major factor hindering the use of ATR-SSFP in clinical routine is the nontriviality in adjusting acquisition parameters, where small adjustments in τ , TR , and RF phase cycling can greatly influence image quality. In this study, we have tested the sequence performance for multiple τ values, while choosing the appropriate phase cycling to position the stop-band at the frequency of hydrogen atoms in fat. The duration of TR_1 was minimized with the constraint that TR_2 had to be sufficiently long to contain the RF pulse. From these options, we selected the settings that were most optimal for abdominal imaging. These settings should allow to introduce ATR-SSFP in the clinic as a simple sequence, for which typical clinical settings, as resolution, SENSE, and FOV, can be adjusted freely as long as τ , TR , and RF phase cycling do not change.

In FS-bSSFP, there is more freedom in the choice of acquisition parameters, and the image quality changes more gradually with adjustments. Adjusting acquisition parameters could improve the quality of FS-bSSFP. For instance, one could increase the number of start-up echoes at the start of each shot to improve the stability of the sequence, decrease the shot length to minimize the recurrence of fat signal, increase the TR to allow for longer RF pulses and thus higher FAs for the same SAR, or increase the resolution by acquiring more k-lines. However, all these adjustments increase scan time. As in our application scan time is limited to a breath-hold, improving one of these parameters is only possible if another one is worsened. For that reason, the full optimization of conventional FS-bSSFP is difficult because it always involves a compromise between parameters that are often conflicting. We believe that changing the settings of FS-bSSFP from the settings used in this study will not significantly improve overall image quality.

The slab excitation profile for (sinc) RF excitation pulses improves with increasing the amount of zero crossings. Increasing the number of pulse zero crossings, without extending its duration or decreasing the FA , leads to higher SAR. To keep SAR low, an increase in the number of zero crossings or FA is often accompanied by the lengthening of the pulse duration, which might collide with the timing constraints of

ATR-SSFP. For that reason, the chosen FA (18°) fell slightly out the range for maximum CNR ($24\text{--}26^\circ$). However, we think that the gain in image quality due to using an RF pulse with a better slab profile, as well as with a FA closer to FA of maximum SNR ($FA = 13\text{--}16^\circ$), outweighs the possible loss in CNR. When considering FS-bSSFP, this problem becomes more severe due to the additional SAR deposition that results from using fat saturation pre-pulses. Here, the FA s in the range from 42 to 47° yield the highest CNR, but in our conventional implementation of FS-bSSFP, the FA was limited to 30° due to SAR constraints. As discussed previously, higher FA s can be obtained, but at a cost of other acquisition parameters and thus poorer image quality.

Contrary to ATR-SSFP, the image quality in FS-bSSFP is degraded as a result of the periodic interruption in the steady-state condition. This interruption results in signal intensity modulations through k-space. This causes image blurring and possibly additional artifacts. The effect of the interruption of k-space can be limited by introducing shorter shot lengths and by adding additional start-up echoes at the start of each shot. However, as discussed previously, due to the limited scan time, this is only possible at the cost of other settings.

There are some limitations associated with this study. Firstly, there was a poor ICC for the scoring of fat suppression. We often observed that, in regions where fat suppression succeeded, fat suppression was stronger for ATR-SSFP (Figs. 6.5 and 6.6). However, in our scan protocol, we focussed shimming on a region around the pancreas. Therefore, for ATR-SSFP fat suppression sometimes failed outside this region. As a result, depending on the preference of the clinician, scoring results may differ. Furthermore, fat saturation in ATR-SSFP may improve when shimming is not focussed on a limited region around the pancreas. Another limitation in our study was that only a subgroup of the patients received hyoscine bromide before the ATR-SSFP and FS-bSSFP scans. We had the impression that administering hyoscine bromide improves the image quality and should be done in future studies. In addition, the highest achievable resolution depends on the acceleration technique that is used, as well as the desired FOV, and, therefore, should be investigated in more detail per application. Here we showed that for our application, a 1.4-mm resolution resulted in robust images whereas artifacts were observed in the 1.2-mm resolution images. Finally, the optimal settings may differ per MRI scanner/vendor, depending on specific implementations. However, for most vendors, shorter TR 2s are not feasible due to the length of the RF pulse and therefore higher τ or other phase cycling schemes are not feasible. In addition, the ideal FA we suggest came from simulations, and therefore are not vendor specific.

One patient failed the breath-hold, and 2 patients did not fit in the FOV in the phase encoding direction. For these patients, image quality was poor. This is reflected

in the fact that when splitting the patients into 2 groups, the images of these 3 patients scored 2.5 (ATR-SSFP) and 1.4 (FS-bSSFP) when averaged over all items scored, whereas the images from the rest scored 2.9 (ATR-SSFP) and 2.1 (FS-bSSFP). Therefore, if the acquisition protocol for ATR-SSFP and FS-bSSFP was adapted for these patients, the scores may increase. As discussed previously, changing settings in ATR-SSFP can greatly influence the overall performance. However, increasing the voxel size or slice thickness to increase FOV or to decrease acquisition time could increase the image quality in the 3 patients mentioned, without greatly affecting the ATR-SSFP performance.

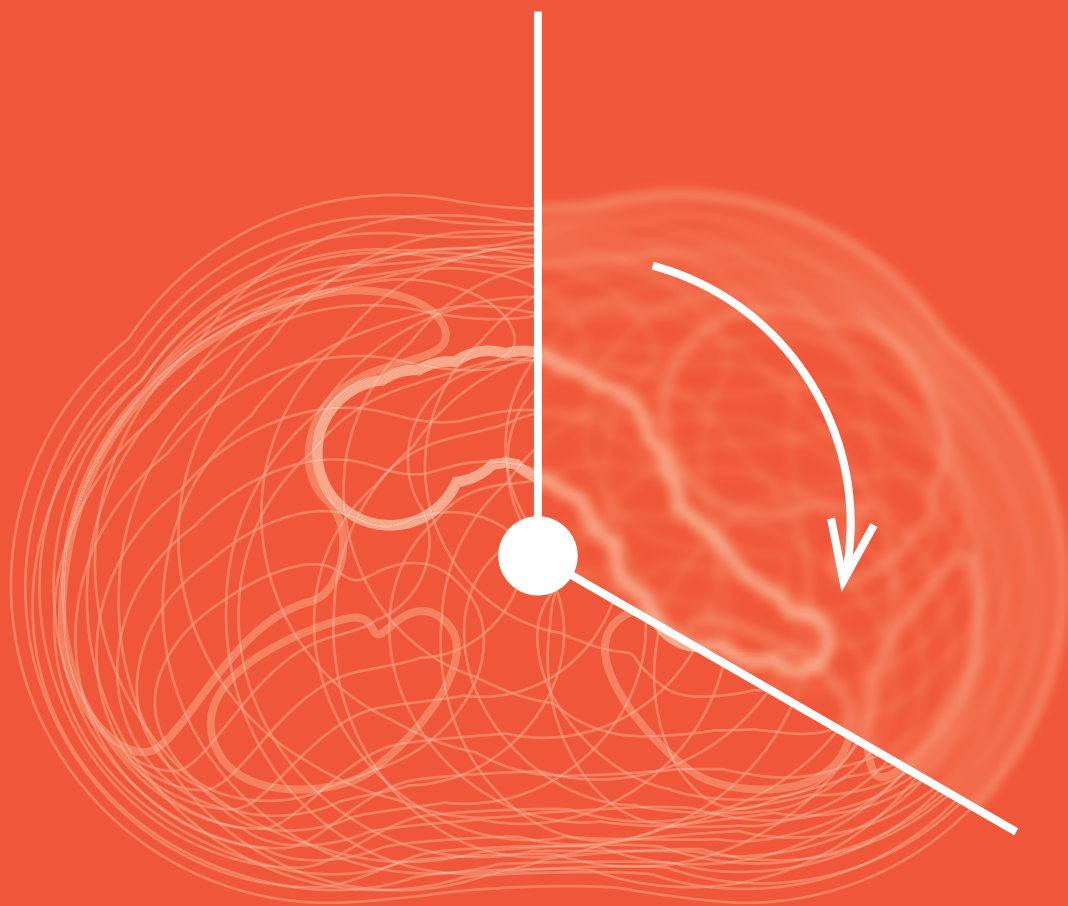
The signal behavior as a function of FA changes for different $T1/T2$ values and settings of τ , RF phase cycling, and field strength [26]. Therefore, it is hard to compare our optimal FA for ATR-SSFP imaging of the abdomen, for which our results show that it appears to be very similar for different organs, with those in the literature. However, we believe that our findings regarding the optimum FA will serve as guidelines for ATR-SSFP of the abdomen, where higher FA s of 30–40° have been traditionally used on 1.5 T [26] and 3 T [3]. Nonabdominal tissues in the body can have different $T1/T2$ values than the abdominal $T1/T2$ values the FA was optimized for in this research. From our simulations, we found that for $\tau = 3$ the FA that renders maximal signal magnitude lays below 30° for $T1/T2 > 4$ and increases to 49° when $T1/T2 = 1$.

In conclusion, the feasible optimal settings for single breath-hold ATR-SSFP imaging of the pancreas at 3 T are $FA = 18^\circ$, $\tau = 3$, RF phase cycling 0°, 180°, 180°, 0° and $TR1/TE = 3.46/1.73$ milliseconds, and isotropic resolution of $1.4 \times 1.4 \times 1.4 \text{ mm}^3$. With these settings, ATR-SSFP imaging is superior to FS-bSSFP, it provides inherently sharper images, and has the potential to be a valuable additional tool to clinical routines in abdominal MRI.

References

1. Miyazaki M, Akahane M. Non-contrast enhanced MR angiography: established techniques. *J. Magn. Reson. Imaging* 2012; 35(1):1–19.
2. Woters PW, Saranathan M, Xu A, Vasanaawala SS. Inversion-recovery-prepared dixon bSSFP: Initial clinical experience with a novel pulse sequence for renal MRA within a breath-hold. *J. Magn. Reson. Imaging* 2012; 35(4):875–881.
3. Gonçalves SI, Ziech MLW, Lamerichs R et al. Optimization of alternating TR-SSFP for fat-suppression in abdominal images at 3T. *Magn. Reson. Med.* 2012; 67(3):595–600.
4. Scheffler K, Lehnhardt S. Principles and applications of balanced SSFP techniques. *Eur. Radiol.* 2003; 13(11):2409–2418.
5. Guglielmo FF, Mitchell DG, O’Kane PL et al. Identifying decreased peristalsis of abnormal small bowel segments in Crohn’s disease using cine MR enterography: the frozen bowel sign. *Abdom. Imaging* 2015; 40(5):1138–1149.
6. Morita S, Ueno E, Masukawa A et al. Defining juxtapapillary diverticulum with 3D segmented trueFISP MRCP: comparison with conventional MRCP sequences with an oral negative contrast agent. *Jpn. J. Radiol.* 2009; 27(10):423–429.
7. Glockner JF, Lee CU. Balanced steady state-free precession (b-SSFP) imaging for MRCP: techniques and applications. *Abdom. Imaging* 2014; 39(6):1309–1322.
8. Jung B, Krombach GA, Günther RW, Buecker A. Is Postcontrast trueFISP Imaging Advantageous? *Invest. Radiol.* 2004; 39(9):517–523.
9. Sandrasegaran K, Lin C, Akisik FM, Tann M. State-of-the-art pancreatic MRI. *Am. J. Roentgenol.* 2010; 195(1):42–53.
10. Ingle RR, Santos JM, Overall WR et al. Self-gated fat-suppressed cardiac cine MRI. *Magn. Reson. Med.* 2015; 73(5):1764–1774.
11. Spuentrup E, Bornert P, Botnar RM et al. Navigator-gated free-breathing three-dimensional balanced fast field echo (TrueFISP) coronary magnetic resonance angiography. *Invest. Radiol.* 2002; 37(11):637–642.
12. Schwab F, Schwarz F, Dietrich O et al. Free breathing real-time cardiac cine imaging with improved spatial resolution at 3 T. *Invest. Radiol.* 2013; 48(3):158–166.
13. Lim RP, Winchester PA, Bruno MT et al. Highly accelerated single breath-hold noncontrast thoracic MRA: evaluation in a clinical population. *Invest. Radiol.* 2013; 48(3):145–151.
14. Ahrar K, Ahrar JU, Javadi S et al. Real-time magnetic resonance imaging-guided cryoablation of small renal tumors at 1.5 T. *Invest. Radiol.* 2013; 48(6):437–444.
15. Saranathan M, Bayram E, Lee CU, Glockner JF. Breath-held MR cholangiopancreatography (MRCP) using a 3D Dixon fat-water separated balanced steady state free precession sequence. *Magn. Reson. Imaging* 2013; 31(8):1263–1270.
16. de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results. *Radiology* 2004; 230(3):652–659.
17. Hargreaves BA, Vasanaawala SS, Nayak KS et al. Fat-suppressed steady-state free precession imaging using phase detection. *Magn. Reson. Med.* 2003; 50(1):210–213.
18. Kornaat PR, Doornbos J, van der Molen AJ et al. Magnetic resonance imaging of knee cartilage using a water selective balanced steady-state free precession sequence. *J. Magn. Reson. Imaging* 2004; 20(5):850–856.
19. Cukur T, Nishimura DG. Multiple repetition time balanced steady-state free precession imaging. *Magn. Reson. Med.* 2009; 62(1):193–204.
20. Vasanaawala SS, Pauly JM, Nishimura DG. Fluctuating equilibrium MRI. *Magn. Reson. Med.* 1999; 42(5):876–883.
21. Vasanaawala SS, Pauly JM, Nishimura DG. Linear combination steady-state free precession MRI. *Magn. Reson. Med.* 2000; 43(1):82–90.

22. Cukur T, Nishimura DG. Fat-water separation with alternating repetition time balanced SSFP. *Magn. Reson. Med.* 2008; 60(2):479–484.
23. Reeder SB, Wen Z, Yu H et al. Multicoil Dixon chemical species separation with an iterative least-squares estimation method. *Magn. Reson. Med.* 2004; 51(1):35–45.
24. Scheffler K, Heid O, Hennig J. Magnetization preparation during the steady state: Fat-saturated 3D TrueFISP. *Magn. Reson. Med.* 2001; 45(6):1075–1080.
25. Hennig J, Speck O, Scheffler K. Optimization of signal behavior in the transition to driven equilibrium in steady-state free precession sequences. *Magn. Reson. Med.* 2002; 48(5):801–809.
26. Leupold J, Hennig J, Scheffler K. Alternating repetition time balanced steady state free precession. *Magn. Reson. Med.* 2006; 55(3):557–565.
27. Dong HZ, Worters PW, Wu HH et al. Noncontrast-enhanced renal angiography using multiple inversion recovery and alternating TR balanced steady-state free precession. *Magn. Reson. Med.* 2013; 70(2):527–536.
28. Al Saleh H, Hernandez L, Lee KS et al. Rapid isotropic resolution cartilage assessment using radial alternating repetition time balanced steady-state free-precession imaging. *J. Magn. Reson. Imaging* 2014; 40(4):796–803.
29. Coolen BF, Poot DHJ, Liem MI et al. Three-dimensional quantitative T1 and T2 mapping of the carotid artery: Sequence design and in vivo feasibility. *Magn. Reson. Med.* 2016; 75(3):1008–1017.
30. Yoshimura K, Kato H, Kuroda M et al. Development of a tissue-equivalent MRI phantom using carrageenan gel. *Magn. Reson. Med.* 2003; 50(5):1011–1017.
31. Look DC, Locker DR. Time saving in measurement of NMR and EPR relaxation times. *Rev. Sci. Instrum.* 1970; 41(2):250–251.
32. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* 1954; 94(3):630–638.
33. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* 1958; 29(8):688–691.
34. Goerner FL, Clarke GD. Measuring signal-to-noise ratio in partially parallel imaging MRI. *Med. Phys.* 2011; 38(9):5049–5057.
35. Gurney-Champion OJ, Lens E, van der Horst A et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med. Phys.* 2015; 42(5):2638–2647.
36. Stanisiz GJ, Odrobina EE, Pun J et al. T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magn. Reson. Med.* 2005; 54(3):507–512.



CHAPTER 7

Minimizing the acquisition time for IVIM model MRI

Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas

Oliver J Gurney-Champion
Martijn Froeling
Remy Klaassen
Jurgen H Runge
Arjan Bel
Hanneke WM van Laarhoven
Jaap Stoker
Aart J Nederveen

Investigative Radiology 2016;51(4):211–220

Abstract

Objective: The aim of this study was to determine the combination of b-values and signal averages for diffusion-weighted image acquisitions that render the minimum acquisition time necessary to obtain values of the intravoxel incoherent motion (IVIM) model parameters *in vivo* in the pancreas or liver with acceptable reproducibility.

Materials and Methods: For 16 volunteers, diffusion-weighted images, with 14 b-values and 9 acquisitions per b-value, were acquired in 2 scan sessions. The IVIM model was fitted to data from lesion-sized regions of interest (ROIs) (1.7 cm³) as well as organ-sized ROIs in the pancreas and liver. By deleting data during analyzes, the IVIM model parameters, D and f , could be determined as a function of the number of b-values as well as the number of measurements per b-value taken along. For the IVIM model parameters, we examined the behavior reproducibility, in the form of the within-subject coefficient of variation (wCV), as a function of the amount of data taken along in the fits. Finally, we determined the minimum acquisition time required as a function of wCV.

Result: For the lesion-sized ROI, the intersession wCVs were 8%/46% and 13%/55% for D/f in the pancreas and liver, respectively, when all data were taken along. For 1.2 times larger wCVs, acquisition in the pancreas could be done in 5:15 minutes using 9 acquisitions per b-value at $b = 0, 30, 50, 65, 100, 375$ and $500 \text{ mm}^{-2}\text{s}$ and for the liver in 2:15 using 9 acquisitions per b-value at $b = 0, 40$ and $500 \text{ mm}^{-2}\text{s}$.

Conclusions: Acquiring 7 b-values in the pancreas and 3 b-values in the liver only decreases the reproducibility by 20% compared with an acquisition with 14 b-values. The understanding of the behavior of reproducibility as a function of b-values and acquisitions per b-values scanned will help researchers select the shortest IVIM protocol.

Introduction

Diffusion-weighted imaging (DWI) is a promising technique for imaging pancreas and liver diffusivity and diffusivity-based lesion identification and characterization in these organs [1-3]. In DWI, the magnetic resonance (MR) signal is made sensitive to the diffusion of water molecules before the acquisition of the signal. In the classical diffusion model, the signal from diffusing water molecules is assumed Gaussian. Therefore, the signal attenuation is modeled monoexponentially, as a function of the apparent diffusion coefficient (ADC) of the tissue, and the amount of diffusion weighting of the signal (b-value).

However, DWI signal also attenuates due to capillary perfusion. Therefore, the intravoxel incoherent motion (IVIM) model for DWI, a biexponential model, was introduced [4]. This model uses three tissue-specific IVIM parameters to describe the signal attenuation as a function of b-value: the diffusion coefficient D , the pseudo-diffusion coefficient D^* , and the perfusion fraction f . In the pancreas and liver, data obtained with b-values higher than $150 \text{ mm}^2\text{s}$ are more predominantly sensitive to diffusion (D), whereas data obtained with b-values lower than $150 \text{ mm}^2\text{s}$ are also sensitive to perfusion ($D + D^*$). Since its introduction, the biexponential signal attenuation was confirmed in multiple studies [5-19] and related to perfusion [20, 21]. The added perfusion parameters of the IVIM model have shown additional value for lesion characterization in the pancreas and liver [5-11, 17, 22, 23]. In addition, they can be used to evaluate organ perfusion and diffusivity and discriminate between healthy liver and liver fibrosis [12-14]. They also enable treatment response monitoring in various other organs [15, 16]. Finally, f shows an excellent histological correlation with histological tumor features in pancreatic tumors [24].

Determining IVIM parameters requires DW images acquired at multiple b-values, resulting in prolonged measurement time compared with the standard monoexponential model. For example, in studies that used respiratory compensation, the mean acquisition time of the pancreas and liver acquisitions was 8:30 minutes (range, 2:24–12:00 minutes; median, 7:55 minutes) [5, 6, 11, 13, 14, 22, 25, 26]. Currently, there is no criterion standard for the choice in b-values at which images are acquired and the number of images per b-value that should be acquired. The lack of a criterion standard could lead to either unnecessarily long acquisitions or biased and unreproducible IVIM model parameter fits. Therefore, to have IVIM acquisitions implemented clinically, more insight should be available on the relation between the b-value selection and the number of averages, and the reproducibility and bias of the fitted IVIM model parameters.

Lemke et al. [27] used Monte-Carlo simulations to determine the ideal b-value distribution for a range of IVIM parameter values. They showed that at least 10

b-values need to be acquired for IVIM model to retrieve the parameters used in the simulated data. However, in simulations, the anatomy is assumed to be static. This situation does not hold in practical pancreatic and liver imaging because these organs are highly mobile due to respiratory, peristaltic, and cardiac motion. This motion can lead to misaligned neighboring slices and misalignments between different volumes belonging to different or repeated b-values. In addition, the echo planar imaging (EPI) readout used in DWI is prone to geometric distortions and susceptibility artifacts [28]. These artifacts often occur in the vicinity of air-tissue boundaries at the edge of the lungs, the stomach or the intestines. Finally, the IVIM model is only an approximation of the actual signal attenuation [4, 29] and the measured signal attenuation is influenced by other, additional factors than the IVIM model accounts for [30, 31]. This may lead a bias in the outcome that depends on the b-values sampled (the equivalence of ADC values that are determined from $b = 0$ and $1000 \text{ mm}^2\text{s}$ being overestimated compared with ADC values that are determined from $b = 150$ and $1000 \text{ mm}^2\text{s}$ in DWI-data of a well-perfused organ) [32].

Together, the effect discussed previously may lead to a different reproducibility of and bias in the IVIM model parameters *in vivo* than expected from simulations that use biexponential data. Therefore, the purpose of this study was to determine the combination of b-values and DW images per b-value that renders the minimum acquisition time necessary to obtain reproducible values of the IVIM parameters without a large bias in the pancreas or liver, using *in vivo* data.

Methods

We included 16 healthy volunteers (8 male, 8 female, mean age 28 years old, range 19–47 years) who were scanned on a 3 T Ingenia (Philips Healthcare, Best, the Netherlands) MR imaging (MRI) scanner. The scanner had a maximum gradient strength of 45 mT/m and a peak slew rate of $200 \text{ mT/m per millisecond}$. Data were acquired with a 16-channel phased-array coil anterior to the volunteer and a 10-channel phased-array coil posterior to the volunteer. All volunteers gave written informed consent. Two MRI sessions (1 hour to 20 days apart; mean, 7 days) were performed. In the first MRI session, we scanned DWI twice to assess intrasession reproducibility. In the second session, we scanned DWI once, which we compared with the first scan of the first session to assess intersession reproducibility. For 2 volunteers, only intersession data were scanned. During each session, we also acquired a 3D Dixon (3 echoes) and a multislice 2D T2-weighted turbo spin echo (TSE) sequence as anatomical references (see Table 1 for acquisition details).

Table 7.1. Sequence parameters for the multi-slice 2D DWI (first column), the 3D Dixon (second column) and the multi-slice 2D T_2 -weighted TSE (last column).

	DWI	Dixon	T2-weighted TSE
FOV (mm ²)	432 × 108 (RL × AP)	400 × 350	400 × 369
acquisition matrix	144 × 36	236 × 208	308 × 230
Slices	21	53	45
Slice thickness/gap (mm)	3.7/0.3	1.7/—	5/—
TR/TE/ Δ TE (ms)	> 2350/44/—	4.4/1.15/0.9	779/80/—
FA (°)	90	10	90
BW (Hz/voxel)	62.5 (phase direction)	1602 (frequency)	548.4 (frequency)
SENSE	1.7 (AP)	2/1.5 (LR/AP)	2 (AP)
Respiratory compensation	Respiratory trigger (navigator)	1 breath-hold	3 breath-holds
Fat saturation	Gradient reversal during slice selection + SPIR	Dixon reconstruction	SPAIR
b-values (mm ⁻² s)	0, 10, 20, 30, 40, 50, 65, 80, 100, 125, 175, 250, 375, 500	—	—

TR of the DWI acquisition was determined by the respiratory trigger interval, but at least 2350 ms.

Abbreviations: FOV = Field of view; RL = right–left; AP = anterior–posterior; TR = repetition time; TE = echo time; FA = flip angle; BW = band width; SPIR = spectral presaturation with inversion recovery; SPAIR = spectrally selective attenuated inversion recovery.

DWI protocol

We acquired all DW images with a single shot EPI readout with axial slice orientation (see Table 7.1 for acquisition details). Diffusion weighting was induced in 3 orthogonal directions, and 3 images were obtained per direction. This resulted in 9 acquisitions per b-value. Each of these acquisitions was acquired at a navigator based expiratory trigger that tracked the liver-lung transition. To avoid intertrigger mismatches between slices of 1 single volume, all slices in the volume were acquired during each trigger. This approach necessitates relatively long acquisition periods (2.35 seconds) compared with common respiratory acquisition times (< 1s). Therefore, to reduce intratrigger respiratory motion, the volunteers were instructed to hold their breath during the typical noise produced by the EPI readout and to breathe freely during the navigator acquisition. This instruction was given before the MRI exam using a recorder sound fragment of the sequence. The bandwidth (BW) per voxel in the phase encoding direction used in literature is on average 21 Hz/voxel (range 9–75 Hz) in the pancreas and liver [1, 5–9, 14, 18, 21, 22, 26, 33, 34]. We used a high BW per voxel (62.5 Hz/voxel)

in order to minimize geometric distortions and susceptibility artifacts. To achieve this, we used parallel imaging as well as a small field of view (FOV) in the phase encoding direction, combined with 2 saturation slabs placed to the posterior and anterior of the FOV.

Image processing

Delineations were done on an averaged $b = 0 \text{ mm}^2\text{s}$ image (averaged over the 9 acquisitions) under the guidance of an averaged $b = 100 \text{ mm}^2\text{s}$, water only reconstruction of the Dixon, and $T2$ -weighted TSE images. The first region of interest (ROI), $\text{ROI}_{\text{large}}$, consisted of the entire pancreas or 2 to 4 slices in the liver. The second ROI, $\text{ROI}_{\text{small}}$, was lesion sized (1.7 cm^3) and spanned 3 slices in the pancreas tail or liver. The $\text{ROI}_{\text{small}}$ s were mostly $3 \times 4 \times 4$ voxel squares located in 3 adjacent slices either in the center of the liver or the tail of the pancreas. However, the pancreas is a small organ and sometimes the squares were deformed to fit in the pancreas. Using anatomical landmarks, such as vessels, ducts, and organ edges, the ROIs were placed at the same location in the different scans for the same volunteer.

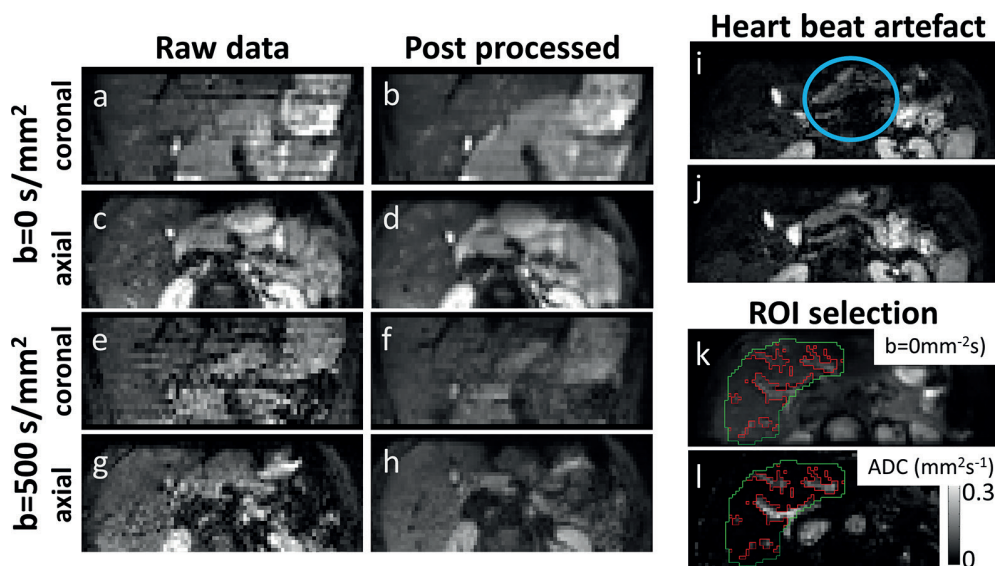


Figure 7.1. Example images from an axial slice (**c, d, g, h**) and coronal multiplanar reformatted slice (**a, b, e, f**) of the data as acquired (**a, c, e, g**) and the data after denoising, cardiac artifact removal, registration, and averaging (**b, d, f, h**) for images from $b = 0 \text{ mm}^2\text{s}$ (**a–d**) and $b = 500 \text{ mm}^2\text{s}$ (**e–h**). An example of signal loss due to the cardiac artifact (**i**, blue circle) compared with an acquisition of the same slice and b -value without the signal loss (**j**). $\text{ROI}_{\text{large}}$ (green) from the liver on the $b = 0 \text{ mm}^2\text{s}$ (**k**) image as well as the ADC image (from $b = 0 \text{ mm}^2\text{s}$ and $10 \text{ mm}^2\text{s}$) (**l**) is shown. The parts selected by the vessel selection algorithm (red) were excluded from $\text{ROI}_{\text{large}}$.

For data processing, a custom-made pipeline was developed in Matlab 2013a (MathWorks, Natick, MA) comprising the following steps. The DW images were denoised using an adaptive optimized non-local means Rician denoising filter as implemented by Manjón et al. [35], with a search radius of 3 voxels and a patch radius of 1 voxel. In some slices, local signal loss due to cardiac motion during diffusion encoding occurred (Fig. 7.1 i-j), which has been reported before [36]. To remove slices containing these cardiac artifacts, we wrote an algorithm that excluded these slices from further analyzes. For each slice and b-value, we calculated the mean and standard deviation (SD) of all voxel values within ROI_{large} of the organ of interest, using all 9 measurements of the same slice. Slices that had a mean intensity in the organ of 2 SDs lower than the mean over all acquisitions of that slice and b-value were removed. On average, the algorithm deleted 7.3% (range between volunteers 3.7%–10.6%) of the slices for the pancreas and 7.9% (range, 4.8%–9.9%) for the liver. To correct for interrespiratory trigger geometric variations and the effect of eddy currents, we registered all DW images group wise using a 4D non-rigid b-spline algorithm based on mutual information using Elastix [37, 38]. By masking, we excluded the slices selected by the cardiac motion algorithm from the cost function of the registration. After registration, the delineations were optimized to correct for errors due to registration.

Per DWI set, the image processing was performed twice, once using the pancreas ROI and once using the liver ROI.

Data fitting

For the IVIM fits, we used a T_1 and T_2 compensated IVIM model [21] with $T_1 = 725$ milliseconds, $T_2 = 43$ milliseconds for the pancreas, $T_1 = 809$ milliseconds, $T_2 = 34$ milliseconds for the liver, as well as $T_1 = 1932$ milliseconds, $T_2 = 275$ milliseconds for blood [39, 40]. As repetition time (TR) was dependent on the respiratory trigger time, which fluctuated each respiratory cycle, we chose a $TR = 5000$ milliseconds for IVIM model fitting.

Signal coming from large blood vessels decays rapidly when diffusion gradients are being applied due to different coherent flow speeds within a voxel. This can cause the data to behave like a triexponential decay [41], rendering it inappropriate for the biexponential IVIM model. In the large ROIs, we segmented and removed such vessels from the ROI_{large} using the DW images by excluding voxels with an ADC value higher than $0.022 \text{ mm}^2\text{s}^{-1}$ on an ADC image created from $b = 0$ and $10 \text{ mm}^{-2}\text{s}$ [7]. The cutoff value of $0.022 \text{ mm}^2\text{s}^{-1}$ was chosen after visually comparing the performance of the algorithm for several cutoff values in the liver, as the vessels in the liver are easy to detect (typical example in Fig. 7.1 k-l). At this value, the algorithm segmented vessels

in the liver that were also visible on the $b = 0 \text{ mm}^2\text{s}$ DW image (Fig. 7.1) as well as the Dixon and T2-weighted images. To visualize the effect of the segmented vessels on the data from the ROIs, we plotted the data, averaged over all volunteers, from the $\text{ROI}_{\text{large}}$ s before and after vessel segmentation. To check how well these data were described the IVIM model, we fitted a biexponential fit [4] and a triexponential fit [41] to it. For biexponential fits to the data from the ROIs with and without vessels we calculated the adjusted R^2 value to determine how well these fits described the data.

All fits were done using the unconstrained Levenberg-Marquardt least squares fit. Unless mentioned otherwise, we fixed the value of D^* to the value obtained from the fits to the averaged data of all $\text{ROI}_{\text{large}}$ s from all volunteers [5-7]. To examine the influence of fixing D^* on D and f , we determined D and f per patient per DWI acquisition, using both the D^* -free and D^* -fixed fitting methods. We compared for significant ($p \leq 0.05$) differences between both methods using a paired t-test for D and f . In addition, to get an idea of fluctuations in D , f , and D^* over this healthy population, we calculated the between-subject coefficient of variations ($b\text{CVs}$, also known as coefficient of variations), using the data from only the first scan of each subject.

Reproducibility of parameter fits

To study the reproducibility and bias of IVIM model parameters as a function of the b -values and number of selected DW images per b -value in the fit, we repeatedly deleted data points before the IVIM model fit.

As a measure for reproducibility, we took the within-subject coefficient of variation ($w\text{CV}$) for D and f : $w\text{CV}_D$ and $w\text{CV}_f$, respectively, for the different fits mentioned previously. The $w\text{CVs}$ were calculated as defined by Barnhart et al. [42] (see Supplemental Digital Content 1, <http://links.lww.com/RLI/A234>, for a short description). We used a Levine test on the percentage difference to check for significant ($p < 0.05$) differences between the intersession and intrasession $w\text{CVs}$ for the full data set (SPSS, version 21, IBM, New York, NY). In addition, for the full data set, Bland-Altman plots [43] were plotted for all ROIs to visualize intersession variations for the pancreas and liver. A 1-sample t-test was done to test for any significant bias ($p \leq 0.05$) of D and f .

As a measure for bias, we took the normalized deviation (ND) of D and f from the D and f values obtained from the full data set. For example, for D , this was defined as

$$ND_D = \frac{D_{\text{mean}} - D_{\text{mean,full}}}{D_{\text{mean,full}}} \times 100\% \quad (7.1)$$

in which D_{mean} was the value found for D averaged over all volunteers using the specific combination of b-values and DW images per b-value and $D_{mean,full}$ was value found for D averaged over all volunteers using the full data set. ND_f was defined similarly.

To visualize the reproducibility and bias as a function of the data taken along, we plotted heat maps of ND and wCV as a function of the number of b-values and the acquisitions per b-value used. This was done for intersession and intrasession, liver and pancreas, ROI_{small} and ROI_{large} and with D^* as a fixed and as a free parameter in the IVIM model. In addition, deleting b-values was done according to three different preselected schemes: a scheme in which mainly low b-values remain in the data set (b-low), a scheme in which mainly high b-values remain (b-high) and a scheme in which b-values are deleted randomly (b-both). In these schemes, b-values were removed in the following order: b-low: 175, 375, 65, 125, 80, 30, 250, 50, 10, 100, 20 and 40 mm^2s , b-high: 20, 40, 10, 65, 175, 80, 30, 100, 375, 50, 250 and 125 mm^2s and b-both: 125, 80, 40, 20, 250, 10, 175, 65, 375, 30, 50 and 100 mm^2s . All schemes kept $b = 0$ and 500 mm^2s . Images per b-value were deleted in a way that the directional spread of the diffusion weighting was as homogeneous as possible.

To investigate the relation between scan time and reproducibility, we selected the combinations of b-values, acquisitions per b-value and b-scheme that had both wCV s lower than 1 to 2 (steps of 0.1) times the wCV s from the full acquisitions. To ensure robustness of these combinations, we then selected the combinations of which the neighbors also satisfied the wCV criterion. From these combinations, we selected the one with the lowest number of acquisitions (b-values \times acquisitions per b-value). Finally, by assuming a respiratory cycle of 12 per minute, the shortest acquisition time was plotted as a function of increasing loss in reproducibility.

As the performance of our image processing may depend on the b-values and the number of selected DW images per b-value, we repeated the image processing for the combination that yielded the fastest acquisition without increasing the wCV s by more than 20% and compared the wCV s and ND s to the case where image processing was done with the full data set. In addition, D - and f -maps were calculated for these settings and compared with the D - and f -maps from the full data set.

Results

Data fitting

Using data from ROIs including large vessels, the IVIM model fit was inaccurate at describing data from low b-value when compared with a triexponential fit (Fig. 7.2).

When vessels were excluded, the IVIM model described the data as accurate as the triexponential fit, both for the pancreas and liver. The R^2 values of the biexponential fits were 0.91 and 0.77 for the pancreas and liver, respectively, when vessels were included and increased to 0.92 and 0.82 when vessels were excluded, which can also be visually appreciated in Fig. 7.2. Therefore, the $\text{ROI}_{\text{large}}$ excluding vessels was selected for further analyzes, which resulted in $\text{ROI}_{\text{large}}$ of 29 cm^3 on average (range, 18–43 cm^3) for the pancreas and 33 (range, 13–59) cm^3 for 2 to 4 slices of the liver. We found $D^* = 70.3 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$ in the pancreas and $D^* = 58.6 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$ in the liver. In $\text{ROI}_{\text{small}}$, the vessels were excluded during delineation.

D and f were calculated for the entire group using the D^* fixed and D^* free IVIM model fits (Table 7.2). There was no significant difference for the fitted D between both fitting methods. However, f was significantly different in all cases (liver, pancreas, $\text{ROI}_{\text{small}}$, and $\text{ROI}_{\text{large}}$). The $b\text{CVs}$ (Table 7.3) were similar for the D^* -free and D^* -fixed fits.

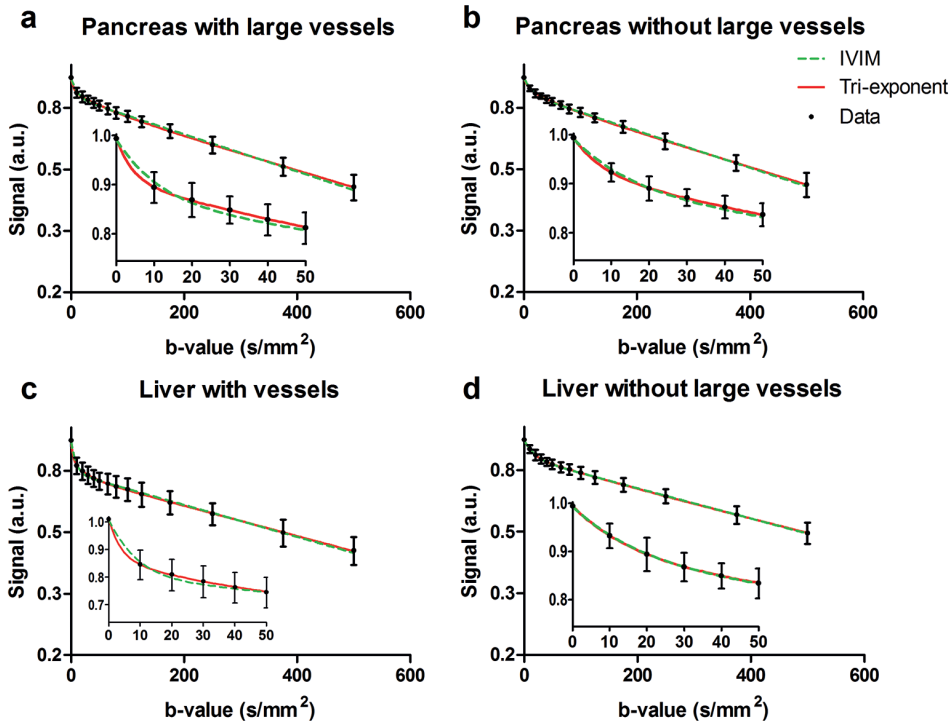


Figure 7.2. Log plots of the data from the signal at different b-values averaged over all 3 measurements from all 16 volunteers of the pancreas (a and b) and the liver (c and d) using the large ROIs with (a and c) and without (b and d) large vessels. The IVIM model (green dashes) and a triexponential decay (red line) were fitted to the data. Subplots show a magnification of the low b-values. The error bars indicate the SD between volunteers of the first scan.

Reproducibility of parameter fits

Both intersession and intrasession wCVs were better for the pancreas than for the liver (Table 7.4). The wCVs were similar to the bCVs. Six out of 8 wCVs mentioned in Table 7.4 were lower for intrasession than intersession, and one was equal. However, none of the differences were significant according to the Levine test. The Bland-Altman plots (Fig. 7.3) from the $\text{ROI}_{\text{small}}$ s for D and f did not show any significant bias.

The intersession and intrasession results on all individual b-schemes, all ROIs and both organs, for fits with a fixed D^* as well as D^* as free model parameter, were calculated (see graphs, Supplemental Digital Content 2, <http://links.lww.com/RLI/A235>, and Supplemental Digital Content 3, <http://links.lww.com/RLI/A236>, which show heat maps for all discussed situations). Generally, both for pancreas and liver, when data from fewer b-values and fewer DW images per b-value were selected, the wCVs and NDs became worse (e.g., Figs. 7.4 and 7.5). In addition, the heat maps changed between the different schemes for deleting b-values (Supplemental Digital Content 2, <http://links.lww.com/RLI/A235>, and Supplemental Digital Content 3, <http://links.lww.com/RLI/A236>).

Table 7.2. Mean values for the population \pm the SD of the value over the population.

	D^* -fixed		D^* -free		
	$D \pm \text{SD}$ ($10^{-3} \text{ mm}^2 \text{ s}^{-1}$)	$f \pm \text{SD}$ (%)	$D \pm \text{SD}$	$f \pm \text{SD}$	$D^* \pm \text{SD}$ ($10^{-3} \text{ mm}^2 \text{ s}^{-1}$)
Pancreas					
ROI _{large}	1.41 \pm 0.17	5.4 \pm 1.1	1.37 \pm 0.13	6.0 \pm 1.2	80 \pm 34
ROI _{small}	1.34 \pm 0.18	6.9 \pm 2.6	1.31 \pm 0.19	7.7 \pm 2.5	1.8 \pm 1.5 $\times 10^2$
Liver					
ROI _{large}	1.12 \pm 0.10	5.1 \pm 1.9	1.10 \pm 0.10	5.5 \pm 1.7	103 \pm 98
ROI _{small}	0.99 \pm 0.12	5.8 \pm 1.5	1.01 \pm 0.10	6.2 \pm 1.6	1.7 \pm 1.1 $\times 10^2$

Table 7.3. Intersession and intrasession bCVs using 14 b-values and 9 measurements per b-value.

	D^* -fixed		D^* -free		
	bCV_D (%)	bCV_f (%)	bCV_D (%)	bCV_f (%)	bCV_{D^*} (%)
Pancreas					
ROI _{large}	13	22	11	36	72
ROI _{small}	15	52	17	48	152
Liver					
ROI _{large}	12	27	15	32	60
ROI _{small}	20	35	20	31	96

Table 7.4. Inter-session and intra-session wCVs using 14 b-values and 9 measurements per b-value.

	Intra		Inter	
	wCV _D (%)	wCV _f (%)	wCV _D (%)	wCV _f (%)
Pancreas				
ROI _{large}	7	19	5	23
ROI _{small}	8	33	8	46
Liver				
ROI _{large}	9	32	10	34
ROI _{small}	12	36	13	55

For the fixed D^* model, the $|ND_D|$ was less than 5% and $|ND_f|$ was less than 15% for most of the combinations, meaning that decreasing the acquisition time did not introduce a large bias. When the number of acquisitions per b-value was 1, 2 or 4, $|ND|$ mostly increased. In these cases, one of the diffusion directions was strongly overrepresented, which was not the case in the reference data set (3 orthogonal acquisitions repeated 3 times). For more than 4 images per b-value, this effect decreased.

D and f from fits with fixed D^* (e.g., Fig. 7.4) were more reproducible (lower wCV) and had less bias (lower $|ND|$) for shorter acquisitions than when D^* was set as free fitting parameter (e.g., Fig. 7.5). In addition, when D^* was set as free fitting parameter, the inter-session wCV of D^* (wCV_{D*}) for the full data set was large: 64% and 134% for ROI_{large} and ROI_{small}, respectively, in the pancreas and 150% and 77% for both ROIs in the liver. The intra-session wCVs for D^* were 61% and 123% for ROI_{large} and ROI_{small}, respectively, in the pancreas and 133% and 88% for both ROIs in the liver.

When wCVs larger than the wCVs found from the full data set were accepted (Table 7.4), IVIM could be acquired in shorter scan times (Fig. 7.6). For the pancreas in ROI_{large} and the liver in ROI_{small}, shorter scan times could be used without decreasing the reproducibility. When D^* was fixed, the scan time could be halved to quartered, without increasing the wCV by more than 20% relative to the full acquisition, for all ROIs. When D^* is set as a free fitting parameter, the decrease in scan time was less pronounced (see graphs, Supplemental Digital Content 3, <http://links.lww.com/RLI/A236>, Fig. 7 U, which is similar as Fig. 7.6 except D^* is set as free fit parameter).

The fastest acquisition that had less than 20% additional error for ROI_{large} was a 3:30 minute scan for the pancreas comprising of 6 images at $b = 0, 30, 50, 65, 100, 375$ and $500 \text{ mm}^{-2}\text{s}$ and a 2:15 minute scan in the liver comprising 9 images at $b = 0, 125$ and $500 \text{ mm}^{-2}\text{s}$. For ROI_{small}, reproducible acquisition was feasible in the pancreas in 5:15 minutes using 9 images at $b = 0, 30, 50, 65, 100, 375$ and $500 \text{ mm}^{-2}\text{s}$ and in the liver in 2:15 minutes using 9 images at $b = 0, 40$ and $500 \text{ mm}^{-2}\text{s}$.

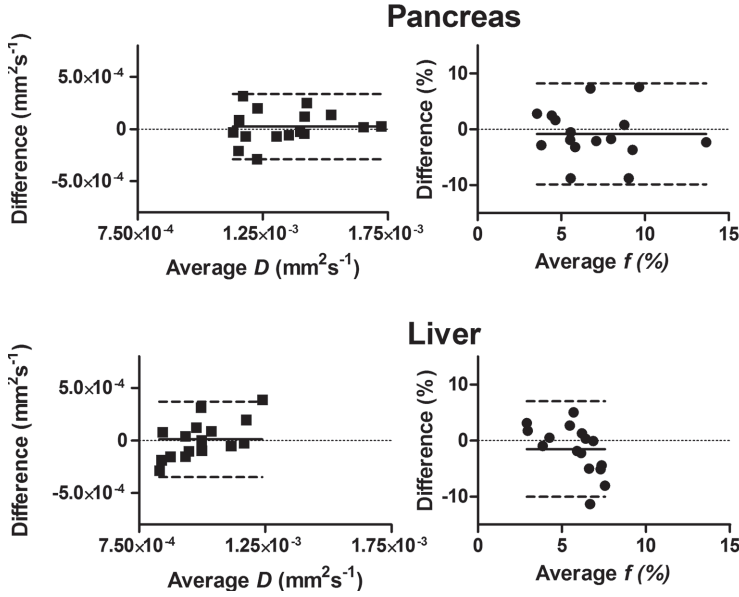


Figure 7.3. Bland-Altman plots for the inter-session variation of D and f from $\text{ROI}_{\text{small}}$ of the pancreas (**top**) and liver (**bottom**) using all b-values and acquisitions per b-value.

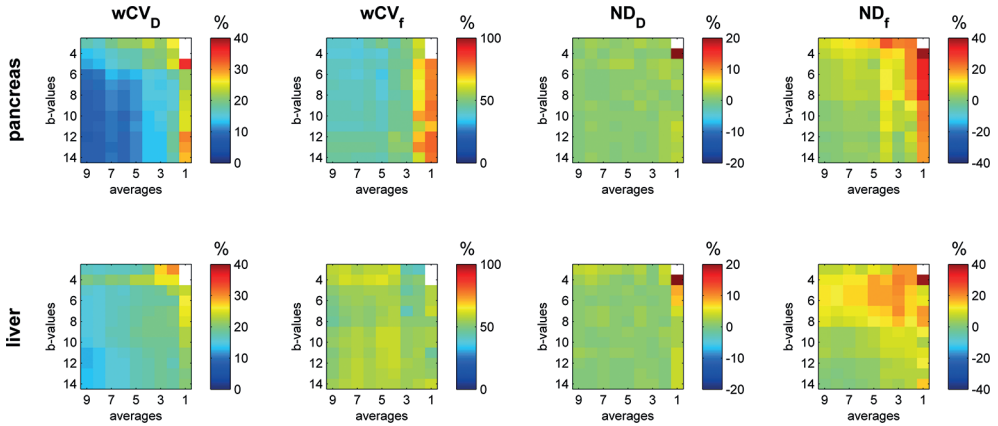


Figure 7.4. From left to right, heat maps of inter-session $w\text{CV}_D$, $w\text{CV}_f$, ND_D and ND_f for $\text{ROI}_{\text{small}}$ in the pancreas (**top row**) and liver (**bottom row**). The vertical axis shows the number of b-values selected, whereas the horizontal axis indicates the number of selected DW images per b-value. The plots show results for deleting b-values using the b-both scheme. White voxels indicate that no values were obtained from the fit, and thus no $w\text{CV}$ or ND could be calculated for this combination.

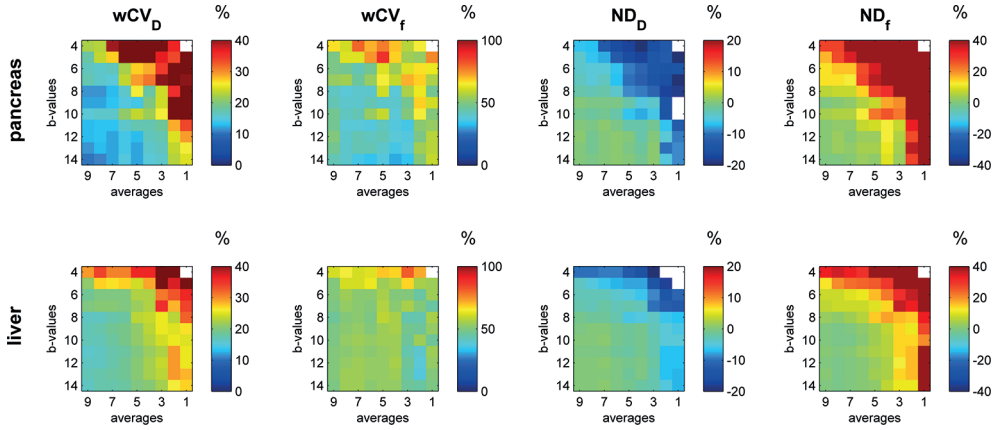


Figure 7.5. Heat maps similar to the ones in Fig. 7.4 from the fits with D^* -free.

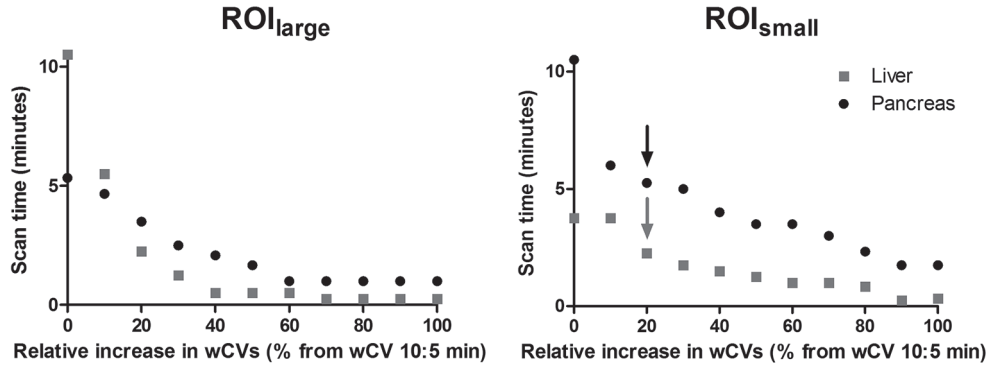


Figure 7.6. Plots of the scan times as a function of additional relative error on intersession wCV for ROI_{large} (left column) and ROI_{small} (right column). The arrows indicate the settings at which Fig. 7.7 was produced.

When image post-processing was done only on the selected data from the fast acquisition of ROI_{small} , the bias and reproducibility changed from $wCV_D = 9\%$, $wCV_f = 44\%$, $ND_D = 0.5\%$, $ND_f = 4.2\%$ to $wCV_D = 10\%$, $wCV_f = 46\%$, $ND_D = 0.7\%$, and $ND_f = 1.5\%$ in the pancreas and from $wCV_D = 14\%$, $wCV_f = 49\%$, $ND_D = 3.8\%$, $ND_f = 5.8\%$ to $wCV_D = 16\%$, $wCV_f = 51\%$, $ND_D = 4.9\%$, and $ND_f = 4.7\%$ in the liver. This indicates that our post-processing also performs well with fewer data. There was little difference in the D - and f -maps reconstructed from the full data set, compared with the limited data set (Fig. 7.7).

Discussion

We used *in vivo* DWI data to determine the minimal number of b-values and acquisitions per b-value necessary to obtain reproducible IVIM parameters in the pancreas and liver with low bias. First, we showed that the IVIM model is only valid in ROIs in which large vessels are excluded. Then, we showed that when fewer b-values and acquisitions per b-value are used, IVIM modeling is more reproducible with D^* fixed than with D^* set as a fitting parameter. Considering this and the poor wCV of D^* , we recommend that D^* is fixed. In addition, we showed that for the pancreas IVIM can be done in 3:30 minutes when changes affecting the entire organ are of interest, and in 5:15 minutes when a specific lesion needs to be measured. Finally, for the liver IVIM can be obtained in 2:15 in both cases.

In this paper, we show that large vessels contribute to an extra attenuation component in the DW-data. This is because signals from flowing blood in large vessels are very sensitive to diffusion gradients, due to the parabolic velocity profile of the blood. Signal thus rapidly decays between $b = 0$ and $10 \text{ mm}^2\text{s}$ for such vessels. Therefore, we recommend that during post-processing, either large vessels are removed from the ROI, or the use of triexponential models is taken into consideration. Alternatively, flow-compensating gradients could be used to overcome this problem at the acquisition level [29].

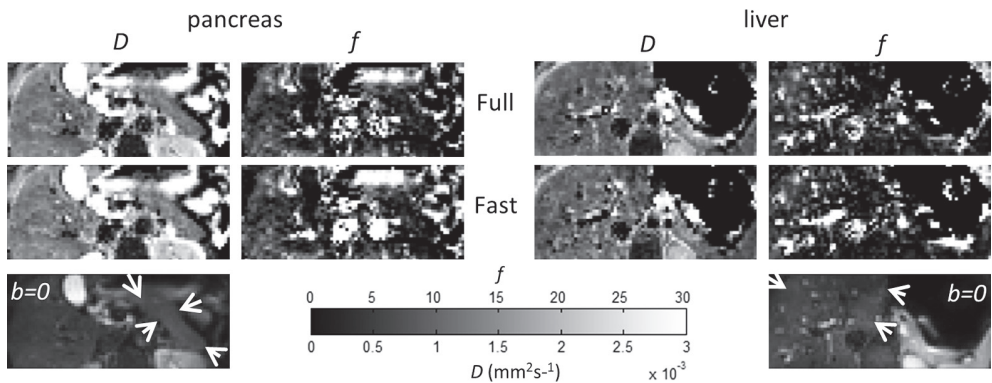


Figure 7.7. An example of the performance of DWI after selecting optimal settings compared with the full scan. The **top row** shows the D - and f -maps obtained with the full set of data. The **middle row** shows the same maps when only the selected sub-set of data is taken along (post-processing performed after deleting b-values). In the bottom row, the $b = 0 \text{ mm}^2\text{s}$ is shown as a reference. The organ of interest (left images, pancreas; right images, liver) is delineated by arrows.

The D^* found by fitting to all data was considerably lower than the mean D^* from the fits to individual patients (Table 7.2). When all data are combined, the signal-to-noise (SNR) ratio becomes higher and D^* can be estimated better than on individual level. In simulations, we indeed found that D^* was overestimated when the SNR decreased (Supplemental Digital Content 4, <http://links.lww.com/RLI/A237>, simulations of fitted parameters as a function of SNR). The signs of the systematic errors found in these simulations were in agreement with the signs of the NDs from the D^* -free fitting, in which D^* and f were overestimated and D was underestimated (Fig. 7.5 and Supplemental Digital Content 3, <http://links.lww.com/RLI/A236>). In addition to a different D^* , there was a small, but significant, difference in f between the D^* -fixed fits and D^* -free fits. This is a result of the different D^* values used in latter.

The D and D^* values reported here for the pancreas and liver agreed well with recent literature: $D = 0.90$ to $1.48 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$ and $D^* = 14.8$ to $59.4 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$ in the pancreas [5, 9, 17, 21, 33] and $D = 1.00$ to $1.17 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$ and $D^* = 13.1$ to $70.6 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$ in the liver [13, 14, 17, 22, 23, 25]. Our fitted values for f in the pancreas and liver are similar to the value found in the pancreas by Lemke et al. [21] ($f = 6.1\%$ – 7.8%), who introduced the T_1 and T_2 correction that we also applied. To compare our values for f with other literature, we recalculated the perfusion fraction, f' , as would be obtained from the classical IVIM model in an acquisition with typical values for TR (4500 milliseconds) and echo time ($TE = 70$ milliseconds). We found an $f' = 16.9\%$ in the pancreas and an $f' = 22.8\%$ in the liver. This is similar to values found in healthy volunteers in other research: $f' = 16\%$ to 40% for pancreas and $f' = 15\%$ to 32% for liver [5, 9, 13, 14, 17, 21–23, 25, 33].

We found that when D^* was fixed, acquisitions at only 7 (pancreas) or 3 (liver) b-values were required for reproducible IVIM modeling with low bias. This is lower than the minimum of 10 b-values required according to previous simulations [27]. However, those simulations also fit D^* , which requires additional data points. When D^* is not fixed, we require 12 b-values for the pancreas and 9 b-values in the liver (see Supplemental Digital Content 3, <http://links.lww.com/RLI/A236>, Fig. 7 U), which is comparable to the simulations [27] and agrees with the observation that at least 2 b-values in the domain $0 < b < 50 \text{ mm}^{-2}\text{s}$ are required to determine an unbiased D^* [44]. Dyvorne et al. [45] showed that in their data, when D^* was fitted only 4 b-values were needed for IVIM in the liver. However, their approach is different as they did not take into account the increase in wCV when fewer b-values are used. We favor this criterion since in longitudinal studies that aim to detect treatment effects or disease progression, the reproducibility directly relates to the statistical power.

In the literature, it is common to measure b-values up to $b = 800$ to $1100 \text{ mm}^{-2}\text{s}$ [5, 9, 13, 14, 17, 22, 23, 25, 33]. In addition, simulations show that high b-values help

to determine D [27]. However, including higher b-values to the protocol increases the TE of the acquisitions of all b-values, and, consequently, leads to SNR decreases for all DW images. This effect was not taken along in previous simulations and cannot be undone by deleting data after the acquisition. Therefore, we chose to limit our maximum b-value to $500 \text{ mm}^2\text{s}$. As our D value was similar to values reported in the literature, we believe limiting the highest b-values did not introduce systematic errors. Including higher b-values would potentially help improve wCV_D , but will worsen wCV_f due to the loss of SNR. As our $wCV_D < wCV_f$ and our wCV_D is comparable to wCV_D s in the liver found in the literature (5%–20%) [13, 23, 34, 45, 46], there is no need for b-values higher than $500 \text{ mm}^2\text{s}$ for IVIM in these organs.

The intrasession wCVs (7%–12% for D and 19%–36% for f , Table 7.4) were not significantly lower than the intersession wCVs, (5%–13% and 23%–55%). This suggests that long term variations are small. Possibly, the difference between inter and intra wCVs becomes even smaller when scanning fasted patients at the same time during the day to minimize the physiological fluctuations.

Typically, for normally distributed acquisition noise, one would expect that reproducibility should increase with increasing scan time. The fact that we could decrease the measurement time substantially at the cost of only a limited loss in intersession wCVs, indicates that other factors that do not decrease with additional acquisitions play a role as well. Such factors could include the presence of image artifacts, the mismatch between the IVIM model and the physiology of liver and pancreas, and short term physiological changes. When patients show pathologies, it may be easier to delineate ROIs and ROI placement could become more reproducible. In these cases, longer acquisitions may still improve the wCVs more substantially than shown in our data.

In the literature of IVIM, multiple measures for repeatability are used [13, 22, 23, 34, 45]. We observed the following measures throughout the literature: the wCV as in this study, the mean of the coefficient of variations calculated per patient, the interscan error and the 95% confidence limits. In this research, we used the method discussed by Barnhart et al. [42], which is in agreement with Bland and Altman [47]. Ignoring the different ways to calculate wCVs, the range of published wCVs for IVIM in the liver are $wCV_f = 10\%$ to 32% and $wCV_D = 5\%$ to 20% when large ROIs ($4\text{--}27 \text{ cm}^3$) are used [13, 23, 34, 45, 46]. While wCVs in our ROI_{large} were in the same range, the $wCV_{f,s}$ in our ROI_{small} were somewhat larger. We expect this is due to our ROI_{small} (1.7 cm^3) being smaller than the typical ROIs used so far.

A limitation of this study is that we did not include data from patients with pancreatic or liver pathologies. wCVs in patients may differ from our wCVs due to various reasons, for example, patients may have biliary duct stents or other metal

devices that disrupt the DWI signal [48], underlining the need of investigating the reproducibility in a separate patient cohort. In addition, D^* may be different in pathologies than in healthy tissue, and the effect of fixing D^* to a reference value from healthy volunteers should be investigated in this case. Furthermore, the bias was deduced using surrogate reference values obtained from the full acquisition. Potentially, there are systematic errors in those reference values and, therefore, the true bias may be higher. However, as $|NDI|$ remains low for the acquisition schemes suggested, we can conclude that the increase in bias from the full data set to our optimized set, is small, and thus scanning longer will not decrease bias. Finally, our protocol only covered 84 mm in cranial-caudal direction, as the protocol was initially implemented for pancreas imaging. The liver is often larger in the cranial-caudal direction and increasing the number of slices, or the slice thickness is needed when the evaluation of the entire liver is desired.

In the literature, there are many different fitting algorithms used for the IVIM model, and the introduced reproducibility and bias in fit parameters will depend on the fitting algorithm [49]. Barbieri et al. [46] showed that a Bayesian probability based algorithm improves the reproducibility and decreases the bias of their IVIM fits. However, the behavior of fitting algorithms may also depend on the number of b-values and acquisitions per b-value used. This falls outside the scope of this research and we restricted ourselves to the Levenberg-Marquardt least squares fit, which is widely available.

To select the ideal settings for reproducible and unbiased IVIM modeling we selected an acquisition that was 20% less reproducible than the full acquisition. Other users may prefer other criteria that can lead to other ideal settings. With the use of the heat maps (see Supplementary Graphs, Supplemental Digital Content 2, <http://links.lww.com/RLI/A235> and Supplemental Digital Content 3, <http://links.lww.com/RLI/A236>), ideal settings for each specific purpose can be deduced.

The bCVs found were similar to the wCVs of the full dataset. This suggests that typical differences between healthy volunteers are lower than the reproducibility of the acquisition protocol. Therefore, for pathologies to be detectable in single subjects with IVIM, the changes will need to be larger than the bCVs and wCVs found in this research, which, for some pathologies, is indeed the case. For instance, when the IVIM model is used for lesion classification in the pancreas, f is 72% to 97% lower in pancreas carcinoma compared with healthy pancreas [5, 9], except in the case of neuroendocrine tumors, in which the IVIM model was not conclusive [9, 17]. In hepatocellular carcinoma, a decrease in f of 64% on average was reported when compared with healthy liver tissue [8]. Besides, it was shown that f changes by 79% between hypo-vascular and hyper-vascular hepatic focal lesions [10]. Moreover,

when compared with healthy liver tissue, D is higher by 5% to 29% in hepatocellular carcinoma, 10% in colorectal metastasis, 53% in hemangioma, and 120% in cysts [8, 22, 23]. These values in majority lay above the b CVs and w CVs from our optimized acquisitions, so these pathologies may be detectable in individual patients using our IVIM protocol. It has also been shown that IVIM-like fits with only three b -values can offer diagnostic value in pancreas cancer patients [50]. Finally, in fibrotic liver, f decreases 23% to 40% when compared with healthy livers, whereas D decreases 7% to 18% [13, 14, 25]. These values are in the order of the b CVs and w CVs from our full acquisition for ROI_{large}. Therefore, even though the IVIM model shows a difference between the healthy volunteer group and the patient group, it will be challenging to classify liver fibrosis for individual patients.

The scan times found in this research are shorter than scan times of typical clinical respiratory compensated scan protocols (mean: 8:30 minutes) [5, 6, 11, 13, 14, 22, 25, 26]. In addition, scanning with the suggested b -values should perform better in terms of reproducibility than other protocols with identical acquisition time, but different b -values. Finally, our w CVs are lower than typical clinically relevant changes. Therefore, we believe that the protocol as presented here, with a known magnitude of bias and reproducibility, should help clinicians introduce DWI for IVIM modeling in the clinic.

Conclusion

Intravoxel incoherent motion model data can be obtained using 7 b -values and 6 acquisitions per b -value when the entire organ is of interest, and with 7 b -values and 9 acquisitions per b -value when measuring in a lesion. For the liver, IVIM can be obtained using 3 b -values and 9 acquisitions per b -value. Using these settings, reproducibility decreased less than 20% from the full acquisition of 14 b -values and 9 acquisitions per b -value. The understanding of the behavior of reproducibility and bias as a function of b -values and acquisitions per b -values scanned will help researchers select the shortest IVIM protocol.

References

1. Matsuki M, Inada Y, Nakai G et al. Diffusion-weighted MR imaging of pancreatic carcinoma. *Abdom. Imaging* 2007; 32(4):481–483.
2. Barral M, Taouli B, Guiu B et al. Diffusion-weighted MR imaging of the pancreas: Current status and recommendations. *Radiology* 2015; 274(1):45–63.
3. Taouli B, Koh DM. Diffusion-weighted MR imaging of the liver. *Radiology* 2010; 254(1):47–66.
4. Le Bihan D, Breton E, Lallemand D et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 1988; 168(2):497–505.
5. Lemke A, Laun FB, Klaub M et al. Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: Comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. *Invest. Radiol.* 2009; 44(12):769–775.
6. Klaub M, Lemke A, Grünberg K et al. Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. *Invest. Radiol.* 2011; 46(1):57–63.
7. Re TJ, Lemke A, Klaub M et al. Enhancing pancreatic adenocarcinoma delineation in diffusion derived intravoxel incoherent motion f-maps through automatic vessel and duct segmentation. *Magn. Reson. Med.* 2011; 66(5):1327–1332.
8. Yamada I, Aung W, Himeno Y et al. Diffusion coefficients in abdominal organs and hepatic lesions: evaluation with intravoxel incoherent motion echo-planar MR imaging. *Radiology* 1999; 210(3):617–623.
9. Kang KM, Lee JM, Yoon JH et al. Intravoxel incoherent motion diffusion-weighted MR imaging for characterization of focal pancreatic lesions. *Radiology* 2014; 270(2):444–453.
10. Yoon JH, Lee JM, Yu MH et al. Evaluation of hepatic focal lesions using diffusion-weighted MR imaging: comparison of apparent diffusion coefficient and intravoxel incoherent motion-derived parameters. *J. Magn. Reson. Imaging* 2014; 39(2):276–285.
11. Qu J, Li X, Qin L et al. Comparison of intra-voxel incoherent motion diffusion magnetic resonance imaging and apparent diffusion coefficient in the evaluation of focal malignant liver masses. *Cancer Transl. Med.* 2015; 1(1):11.
12. Ichikawa S, Motosugi U, Morisaka H et al. MRI-based staging of hepatic fibrosis: Comparison of intravoxel incoherent motion diffusion-weighted imaging with magnetic resonance elastography. *J. Magn. Reson. Imaging* 2015; 42(1):204–210.
13. Patel J, Sigmund EE, Rusinek H et al. Diagnosis of cirrhosis with intravoxel incoherent motion diffusion MRI and dynamic contrast-enhanced MRI alone and in combination: preliminary experience. *J. Magn. Reson. Imaging* 2010; 31(3):589–600.
14. Dyvorne H a, Galea N, Nevers T et al. 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* 2013; 266(3):920–929.
15. Gaeta M, Benedetto C, Minutoli F et al. Use of diffusion-weighted, intravoxel incoherent motion, and dynamic contrast-enhanced MR imaging in the assessment of response to radiotherapy of lytic bone metastases from breast cancer. *Acad. Radiol.* 2014; 21(10):1286–1293.
16. Koh DM. Science to practice: Can intravoxel incoherent motion diffusion-weighted MR imaging be used to assess tumor response to antivascular drugs? *Radiology* 2014; 272(2):307–308.
17. Hwang EJ, Lee JM, Yoon JH et al. Intravoxel incoherent motion diffusion-weighted imaging of pancreatic neuroendocrine tumors: prediction of the histologic grade using pure diffusion coefficient and tumor size. *Invest. Radiol.* 2014; 49(6):396–402.
18. Regini F, Colagrande S, Mazzoni LN et al. Assessment of liver perfusion by intravoxel incoherent motion (IVIM) magnetic resonance – diffusion-weighted imaging: Correlation with phase-contrast portal venous flow measurements. *J. Comput. Assist. Tomogr.* 2015; 39(3):365–372.

19. Doblas S, Wagner M, Leitaos HS et al. Determination of malignancy and characterization of hepatic tumor type with diffusion-weighted magnetic resonance imaging: Comparison of apparent diffusion coefficient and intravoxel incoherent motion-derived measurements. *Invest. Radiol.* 2013; 48(10):722–728.
20. Heusch P, Wittsack HJ, Heusner T et al. Correlation of biexponential diffusion parameters with arterial spin-labeling perfusion MRI: Results in transplanted kidneys. *Invest. Radiol.* 2013; 48(3):140–144.
21. Lemke A, Laun FB, Simon D et al. An in vivo verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. *Magn. Reson. Med.* 2010; 64(6):1580–1585.
22. Andreou A, Koh DM, Collins DJ et al. Measurement reproducibility of perfusion fraction and pseudodiffusion coefficient derived by intravoxel incoherent motion diffusion-weighted MR imaging in normal liver and metastases. *Eur. Radiol.* 2013; 23(2):428–434.
23. Kakite S, Dyvorne H, Besa C et al. Hepatocellular carcinoma: Short-term reproducibility of apparent diffusion coefficient and intravoxel incoherent motion parameters at 3.0T. *J. Magn. Reson. Imaging* 2015; 41(1):149–156.
24. Klauf M, Mayer P, Bergmann F et al. Correlation of histological vessel characteristics and diffusion-weighted imaging intravoxel incoherent motion derived parameters in pancreatic ductal adenocarcinomas and pancreatic neuroendocrine tumors. *Invest. Radiol.* 2015; 50(11):792–797.
25. Lu P-X, Huang H, Yuan J et al. Decreases in molecular diffusion, perfusion fraction and perfusion-related diffusion in fibrotic livers: a prospective clinical intravoxel incoherent motion MR imaging study. *PLoS One* 2014; 9(12):e113846.
26. Klauf M, Gaida MM, Lemke A et al. Fibrosis and pancreatic lesions: counterintuitive behavior of the diffusion imaging-derived structural diffusion coefficient D. *Invest. Radiol.* 2013; 48(3):129–133.
27. Lemke A, Stieltjes B, Schad LR, Laun FB. Toward an optimal distribution of b values for intravoxel incoherent motion imaging. *Magn. Reson. Imaging* 2011; 29(6):766–776.
28. Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. *J. Magn. Reson. Imaging* 2006; 24(3):478–488.
29. Wetscherek A, Stieltjes B, Laun FB. Flow-compensated intravoxel incoherent motion diffusion imaging. *Magn. Reson. Med.* 2015; 74(2):410–419.
30. Yablonskiy DA, Bretthorst GL, Ackerman JJH. Statistical model for diffusion attenuated MR signal. *Magn. Reson. Med.* 2003; 50(4):664–669.
31. Bennett KM, Schmainda KM, Bennett RT et al. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* 2003; 50(4):727–734.
32. Zhang JL, Sigmund EE, Chandarana H et al. Variability of renal apparent diffusion coefficients: limitations of the monoexponential model for diffusion quantification. *Radiology* 2010; 254(3):783–792.
33. Ma C, Liu L, Li YJ et al. Intravoxel incoherent motion MRI of the healthy pancreas: Monoexponential and biexponential apparent diffusion parameters of the normal head, body and tail. *J. Magn. Reson. Imaging* 2015; 41(5):1236–1241.
34. Lee Y, Lee SS, Kim N et al. Intravoxel incoherent motion diffusion-weighted MR imaging of the liver: effect of triggering methods on regional variability and measurement repeatability of quantitative parameters. *Radiology* 2015; 274(2):405–15.
35. Manjón J V, Coupé P, Martí-Bonmatí L et al. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J. Magn. Reson. Imaging* 2010; 31(1):192–203.
36. Kwee TC, Takahara T, Niwa T et al. Influence of cardiac motion on diffusion-weighted magnetic resonance imaging of the liver. *MAGMA* 2009; 22(5):319–325.
37. Klein S, Staring M, Murphy K et al. Elastix: A toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* 2010; 29(1):196–205.

38. Huizinga W, Poot DHJ, Guyader J-M et al. Non-rigid groupwise image registration for motion compensation in quantitative MRI. *Work. Biomed. Image Regist. LNCS* vol. 8545, 2014:184–193.
39. de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results. *Radiology* 2004; 230(3):652–659.
40. Stanisiz GJ, Odrobina EE, Pun J et al. T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magn. Reson. Med.* 2005; 54(3):507–512.
41. Cercueil JP, Petit JM, Nougaret S et al. Intravoxel incoherent motion diffusion-weighted imaging in the liver: comparison of mono-, bi- and tri-exponential modelling at 3.0-T. *Eur. Radiol.* 2015; 25(6):1541–1550.
42. Barnhart HX, Barboriak DP. Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl. Oncol.* 2009; 2(4):231–235.
43. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1(8476):307–310.
44. Cohen AD, Schieke MC, Hohenwarter MD, Schmainda KM. The effect of low b-values on the intravoxel incoherent motion derived pseudodiffusion parameter in liver. *Magn. Reson. Med.* 2015; 73(1):306–311.
45. Dyvorne H, Jajamovich G, Kakite S et al. Intravoxel incoherent motion diffusion imaging of the liver: optimal b-value subsampling and impact on parameter precision and reproducibility. *Eur. J. Radiol.* 2014; 83(12):2109–2113.
46. Barbieri S, Donati OF, Froehlich JM, Thoeny HC. Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs. *Magn. Reson. Med.* 2016; 75(5):2175–2184.
47. Bland JM, Altman DG. Statistics Notes: Measurement error. *BMJ Br. Med. J.* 1996; 313(7059):744.
48. Gurney-Champion OJ, Lens E, van der Horst A et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med. Phys.* 2015; 42(5):2638–2647.
49. Wurnig MC, Donati OF, Ulbrich E et al. Systematic analysis of the intravoxel incoherent motion threshold separating perfusion and diffusion effects: Proposal of a standardized algorithm. *Magn. Reson. Med.* 2015; 74(5):1414–1422.
50. Concia M, Sprinkart AM, Penner A-H et al. Diffusion-weighted magnetic resonance imaging of the pancreas: Diagnostic benefit from an intravoxel incoherent motion model-based 3 b-value analysis. *Invest. Radiol.* 2014; 49(2):93–100.

Supplemental Digital Content 1

wCVs

The wCVs were calculated as defined by Barnhart et al. (2009), but as percentage: the within subject standard deviation divided by the mean parameter value multiplied by 100%. The mean parameter value is averaged both over subjects and measurements. The within subject standard deviation was calculated as the root of the within subject mean of squares. The within subject mean of squares, in our situation, was the mean (over the volunteers) of the squared difference of the repeated measure divided by two.

Supplemental Digital Content 2

Biexponential model, fixed D^*

Intrasession

The mean parameter values for intrasession (Table 7.A), and thus used as reference values for ND , where similar to those for intersession (Table 7.2).

Table 7.A. Average value found for D , f and D^* over all volunteers using the full data set from the intrasession scans with D^* -fixed and D^* -free.

	D^* -fixed		D^* -free		
	$D \pm SD$ ($10^{-3} \text{ mm}^2 \text{ s}^{-1}$)	$f \pm SD$ (%)	$D \pm SD$	$f \pm SD$	$D^* \pm SD$ ($10^{-3} \text{ mm}^2 \text{ s}^{-1}$)
Pancreas					
ROI _{large}	1.40 ± 0.17	5.1 ± 1.2	1.33 ± 0.15	6.1 ± 1.4	69.9 ± 32
ROI _{small}	1.32 ± 0.16	5.9 ± 3.0	1.23 ± 0.14	7.3 ± 3.3	112.9 ± 126
Liver					
ROI _{large}	1.12 ± 0.09	4.7 ± 1.6	1.12 ± 0.09	5.2 ± 1.9	109.5 ± 95
ROI _{small}	0.10 ± 0.14	5.0 ± 1.4	1.04 ± 0.15	5.4 ± 1.2	212.1 ± 151

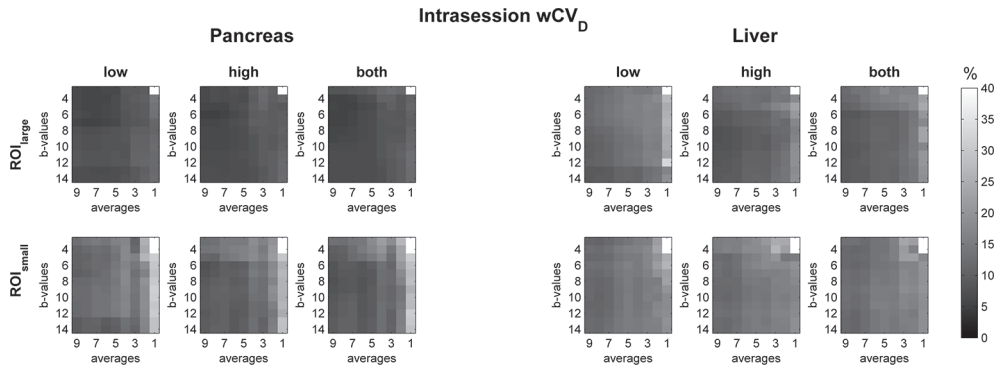


Figure 7.A. Heat maps of the intrasection wCV_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. Heat maps are shown for the pancreas (left) and the liver (right) for several under sampling schemes per organ (low, high and random). Also, the wCV is plotted for ROI_{large} (top row), and ROI_{small} (bottom row). White voxels indicate that no value was obtained for this combination of b-values and averages. Color version is available at <http://links.lww.com/RLI/A235>.

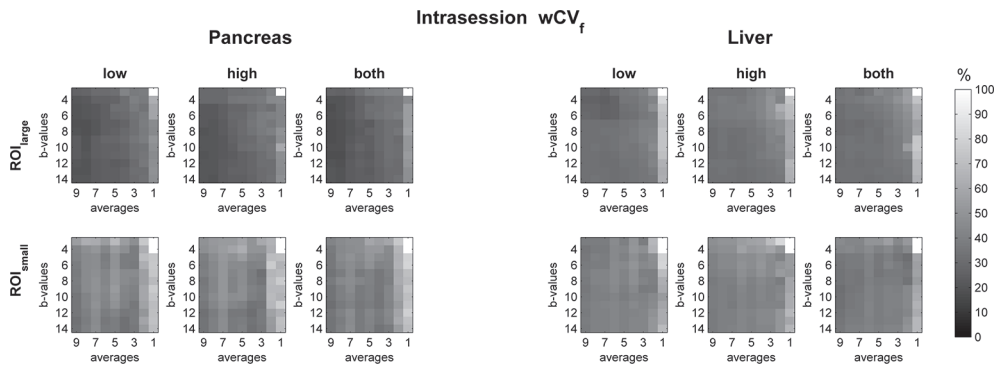


Figure 7.B. Heat maps of the intrasection wCV_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A235>.

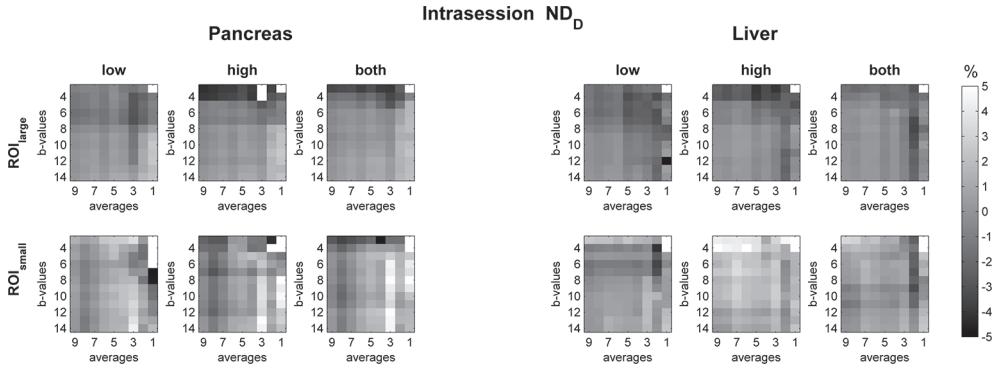


Figure 7.C. Heat maps of the intrasection ND_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A235>.

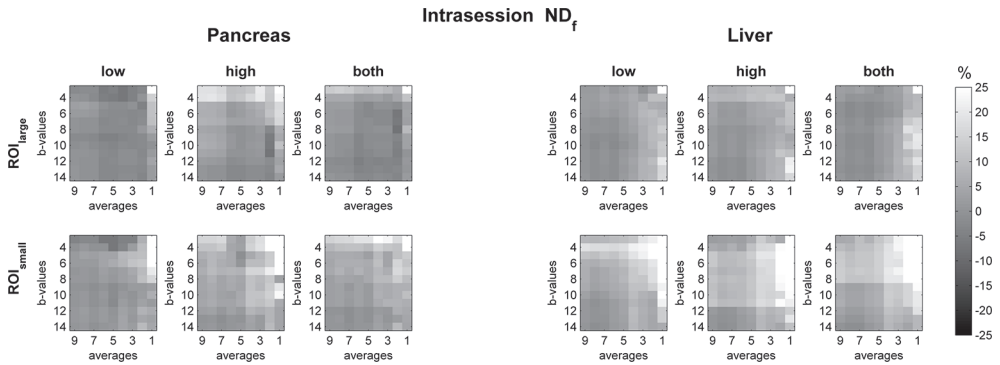


Figure 7.D. Heat maps of the intrasection ND_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A235>.

Interession

The interession graphs for wCVs and NDs (Fig. 7.E-H) are similar to the intrasession graphs. The equivalent of Table 7.A for interession is found as Table 7.1 in the main manuscript.

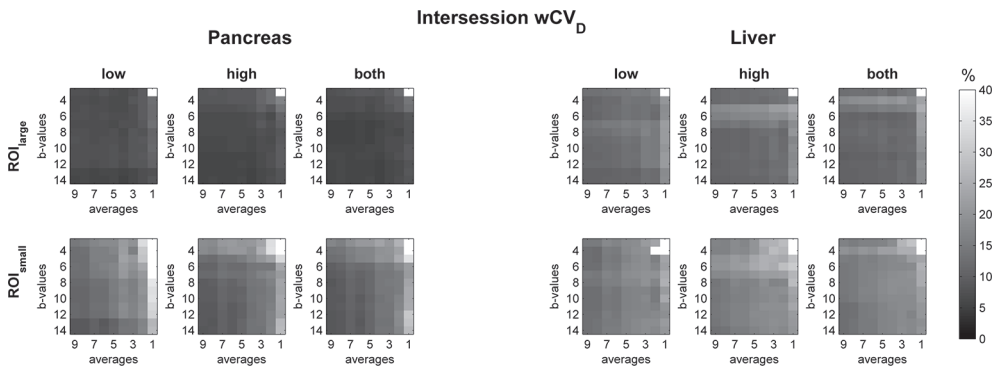


Figure 7.E. Heat maps of the Interession wCV_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A235>.

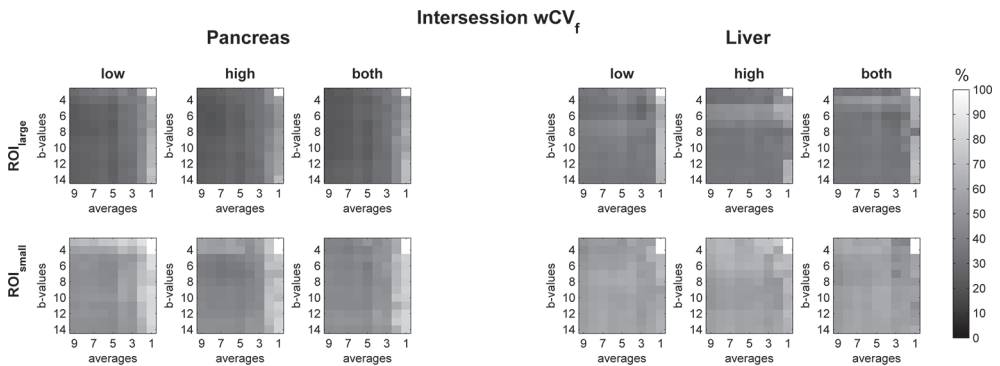


Figure 7.F. Heat maps of the Interession wCV_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A235>.

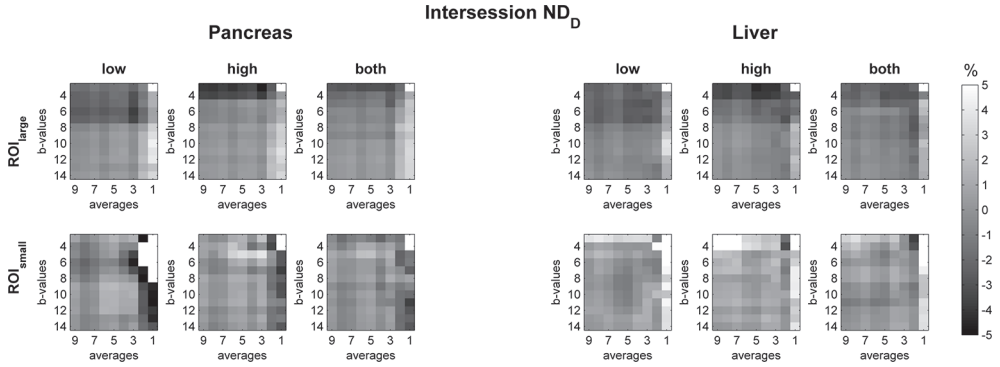


Figure 7.G. Heat maps of the Interession ND_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RL/A235>.

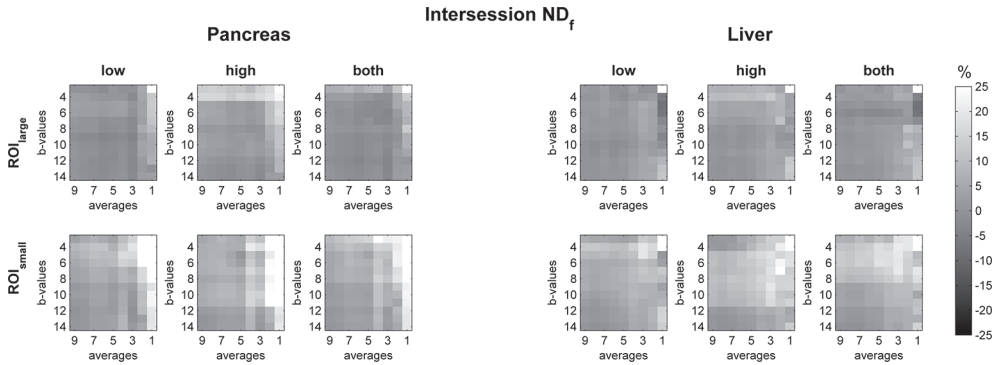


Figure 7.H. Heat maps of the Interession ND_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RL/A235>.

Supplemental Digital Content 3

Biexponential model, free D^*

Intrasection

When D^* was set as fitting parameter, the fitted value for D and f , averaged over all volunteers, where similar to the fixed D^* model when all b-values and measurements were taken along (Table 7.A-B). However, wCVs and NDs became worse, especially when data is deleted. Therefore, when precise IVIM model fitting with high reproducibility is desired with D^* as fitting variable, longer acquisition times are required, in which more b-values and measurements per average are acquired than when D^* is fixed.

Table 7.B. Inter- and intrasection coefficients of variation using 14 b-values and 9 measurements per b-value.

	Intra			Inter		
	wCV _D (%)	wCV _f (%)	wCV _{D*} (%)	wCV _D (%)	wCV _f (%)	wCV _{D*} (%)
Pancreas						
ROI _{large}	8	33	61	7	31	64
ROI _{small}	14	40	123	11	29	134
Liver						
ROI _{large}	10	45	133	12	47	150
ROI _{small}	16	51	88	13	34	77

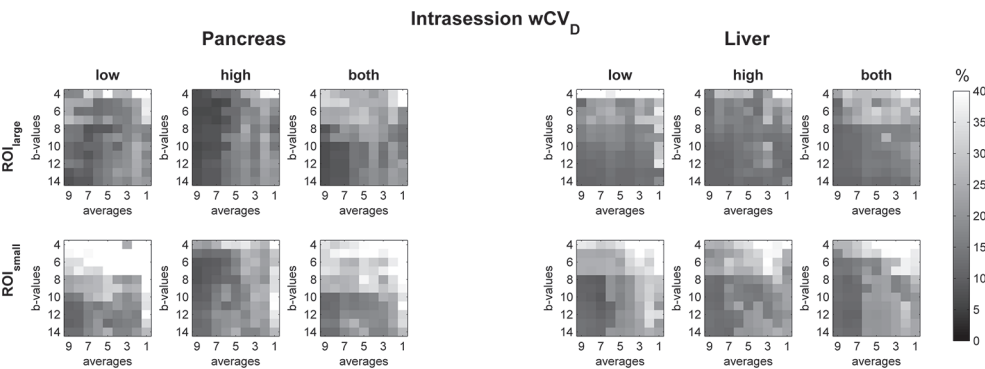


Figure 7.I. Heat maps of the intrasection wCV_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A236>.

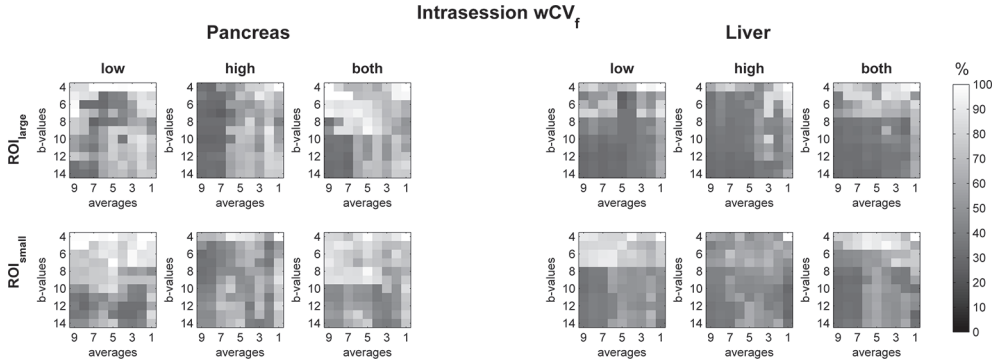


Figure 7.J. Heat maps of the intrasession wCV_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A236>.

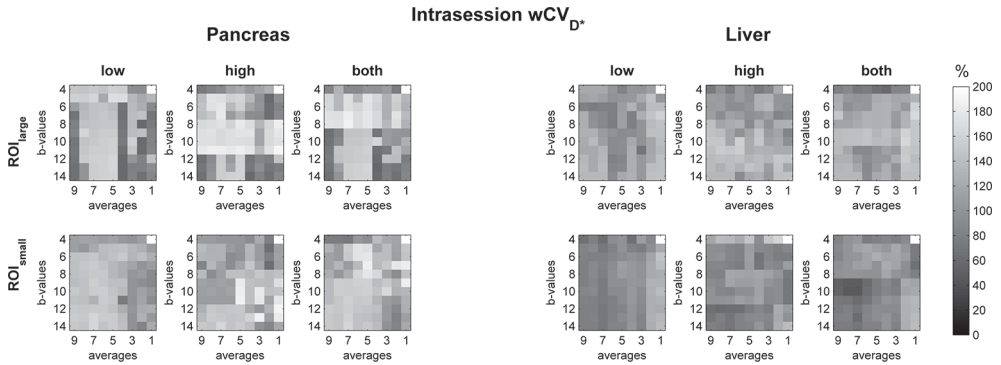


Figure 7.K. Heat maps of the intrasession wCV_{D^*} plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A236>.

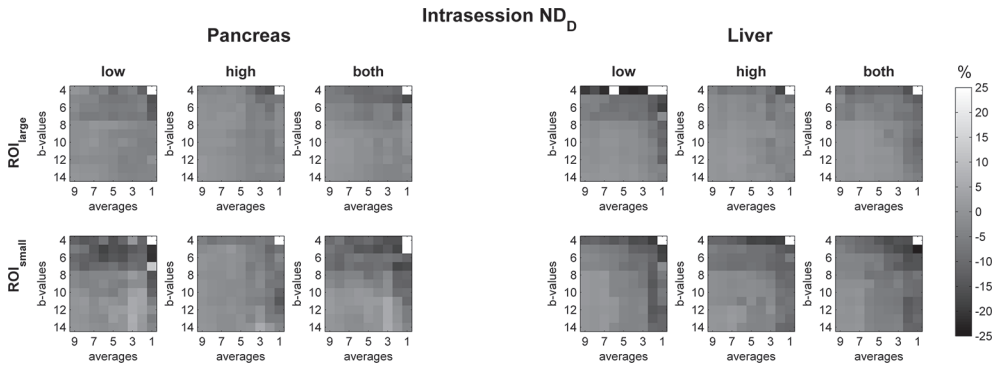


Figure 7.L. Heat maps of the intrasection ND_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs C and G. Color version is available at <http://links.lww.com/RLI/A236>.

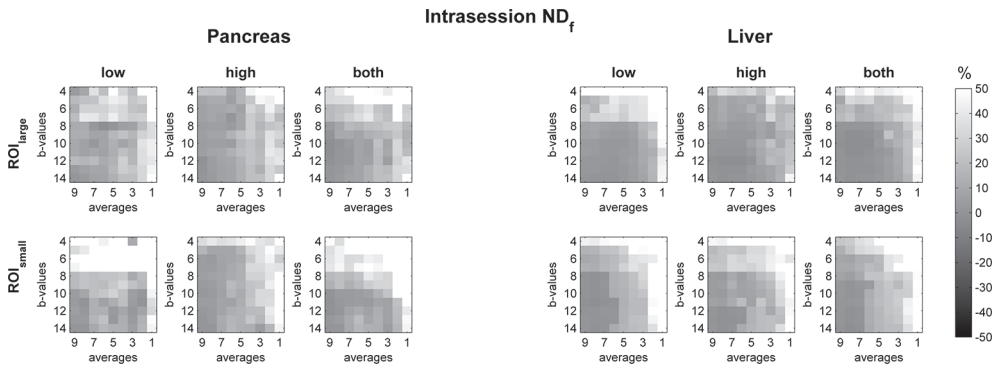


Figure 7.M. Heat maps of the intrasection ND_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs D and H. Color version is available at <http://links.lww.com/RLI/A236>.

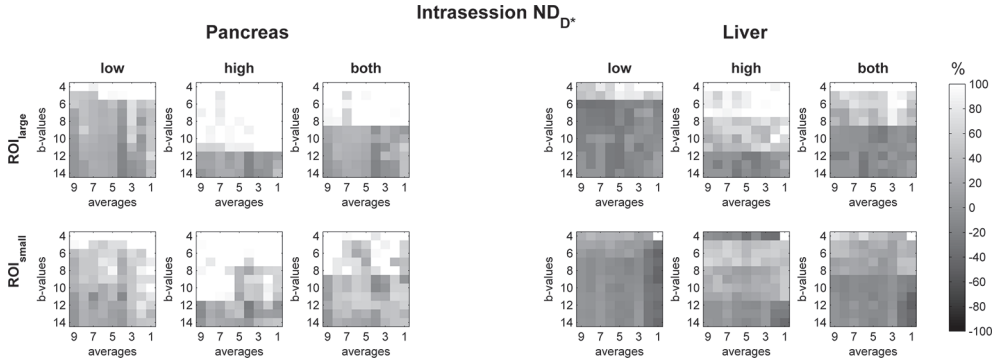


Figure 7.N. Heat maps of the intrasession ND_{D^*} plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RL/A236>.

Intersession

Intersession wCV_s and ND_s show similar behavior as the intrasession versions (Fig. 7.O-T). When 20% larger wCV_s are allowed, precise IVIM model fits with high reproducibility for ROI_{large} can be done in 8 minutes for the pancreas using 12 b-values and 8 acquisitions per b-value and 5:30 minutes for the liver using 9 b-values and 6 acquisitions per b-value. For ROI_{small} it can be done in 8 minutes for the pancreas using 12 b-values and 8 acquisitions per b-value and 5:15 minutes for the liver using 9 b-values and 7 acquisitions per b-value (Fig. 7.U).

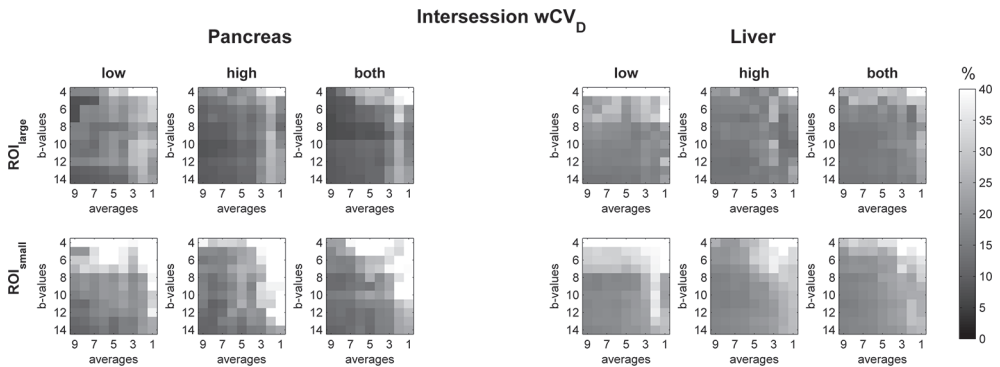


Figure 7.O. Heat maps of the Intersession wCV_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RL/A236>.

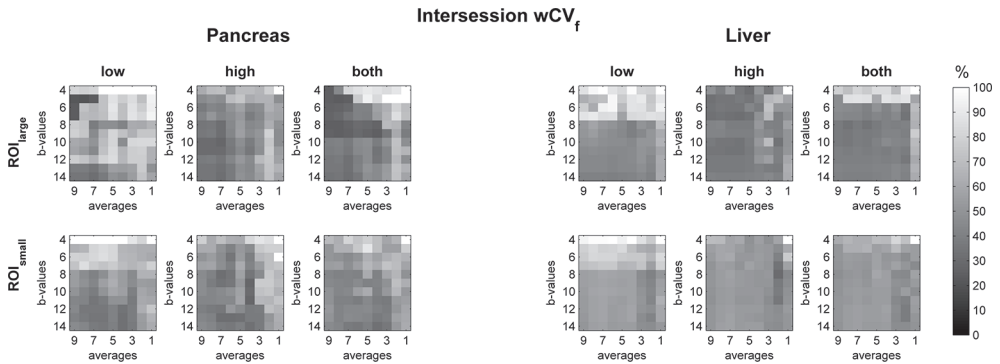


Figure 7.P. Heat maps of the Interession wCV_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A236>.

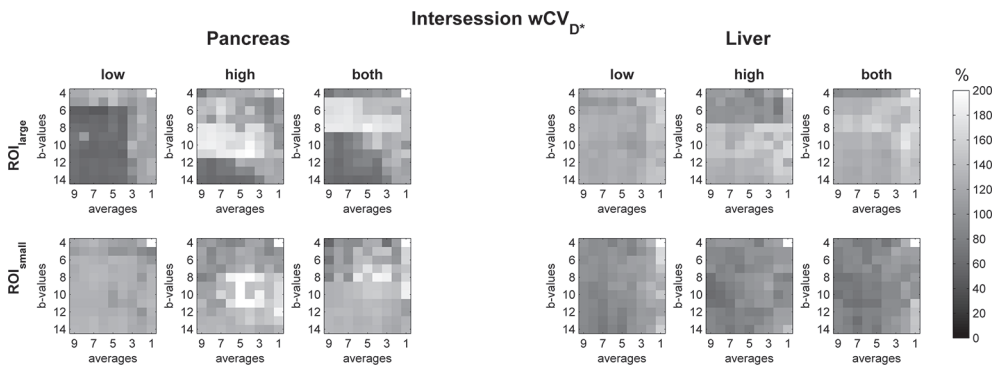


Figure 7.Q. Heat maps of the Interession wCV_{D^*} plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RLI/A236>.

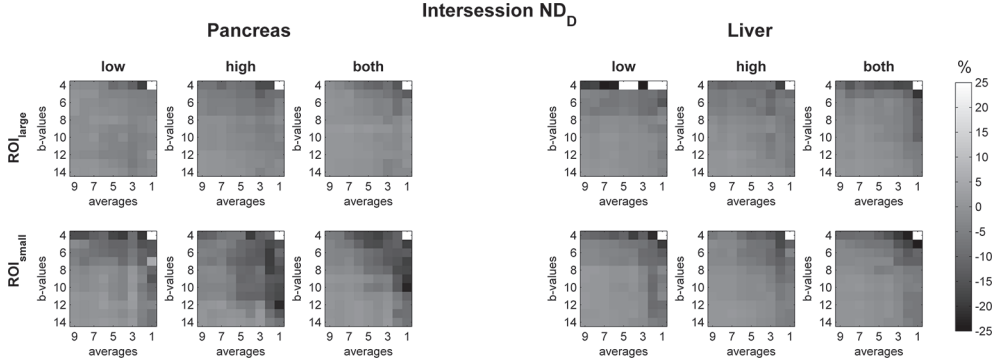


Figure 7.R. Heat maps of the Interession ND_D plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs 7.C and 7.G. Color version is available at <http://links.lww.com/RLI/A236>.

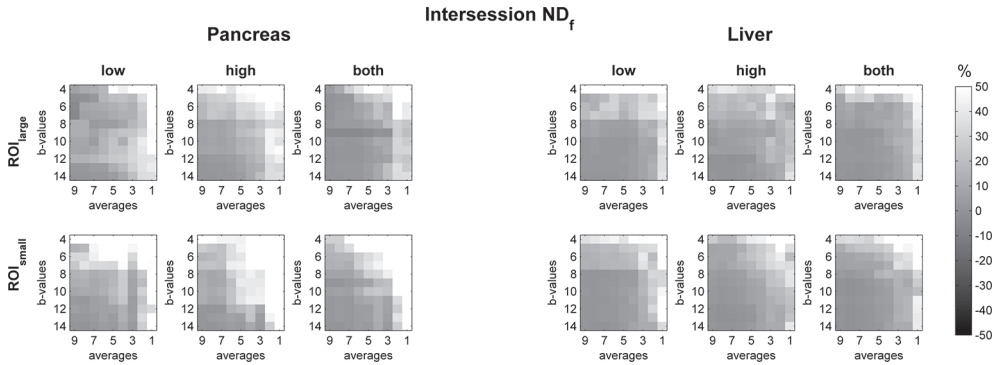


Figure 7.S. Heat maps of the Interession ND_f plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D^* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs 7.D and 7.H. Color version is available at <http://links.lww.com/RLI/A236>.

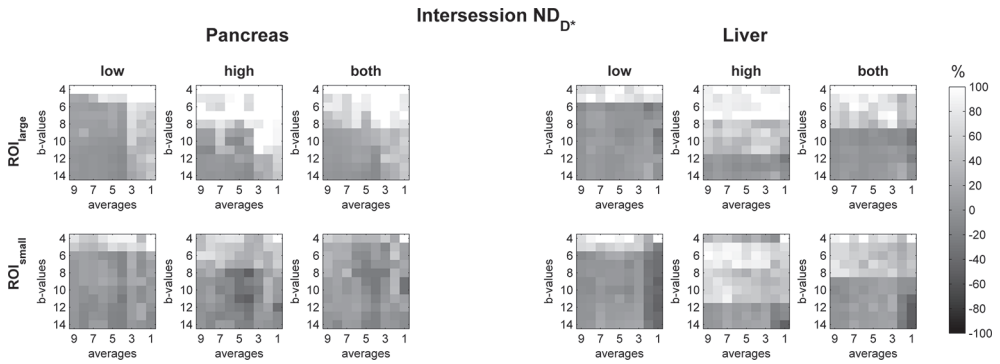


Figure 7.T. Heat maps of the Interession ND_{D^*} plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at <http://links.lww.com/RL/A236>.

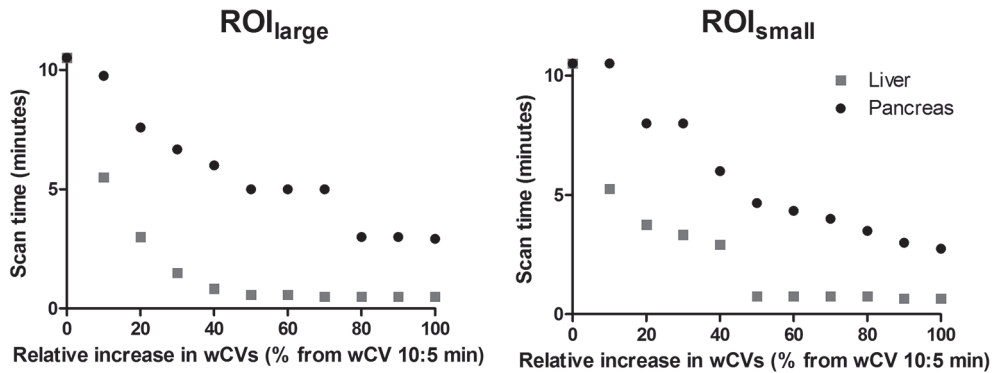


Figure 7.U. Plots of the scan times as function of additional relative error on interession wCV using the ROI_{large} (left) and ROI_{small} (right).

Supplemental Digital Content 4

Simulations

To study the behavior of the fit parameters as a function SNR, we used simulations in Mathematica (Wolfram Research, Champaign, IL). We simulated typical IVIM-data using all b-values from this research and the parameters found from the fit to all volunteer data (pancreas, ROI_{large}). Gaussian noise was added to the data for several SNR values (range 2–100). Then, we fitted the IVIM model to the data. This procedure was repeated 1000 times and the mean IVIM model parameters per SNR level were calculated and plotted as a function of SNR (Fig. 7.V). It is clear that systematic errors are introduced at low SNR.

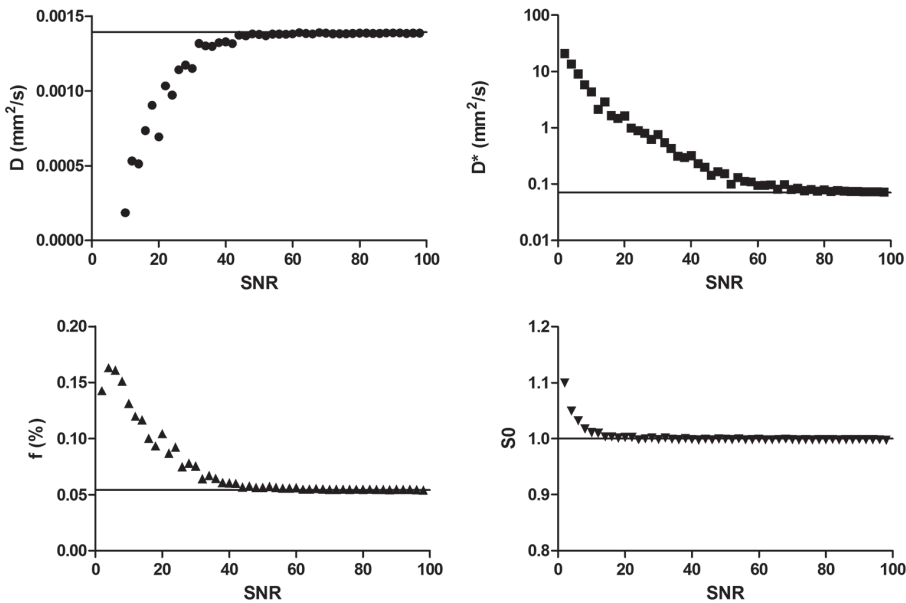
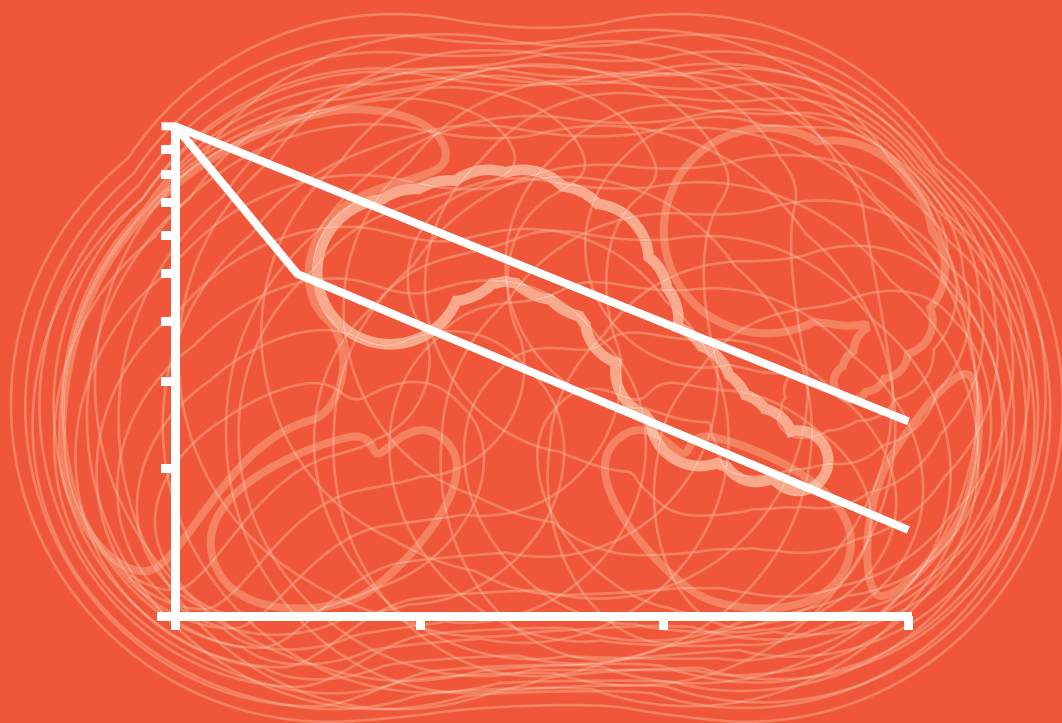


Figure 7.V. IVIM model parameter values from fits on simulated data, as a function of SNR. The stripe indicates the reference value used in the simulation. Note that D^* is plotted on a logarithmic scale.



CHAPTER 8

Best fits for DWI

How best to fit diffusion-weighted magnetic resonance imaging data of pancreatic cancer patients

Oliver J Gurney-Champion
Remy Klaassen
Martijn Froeling
Sebastiano Barbieri
Jaap Stoker
Marc RW Engelbrecht
Johanna W Wilmink
Marc G Besselink
Arjan Bel
Hanneke WM van Laarhoven
Aart J Nederveen

Submitted

Abstract

Purpose: To compare the performance of six intravoxel incoherent motion (IVIM) model fitting algorithms and seven diffusion-weighted imaging (DWI) models for pancreatic cancer imaging.

Methods: The local medical ethics committee approved this prospective study. Patients gave written informed consent. Between October 2014 and March 2016, diffusion-weighted imaging (DWI) data were acquired in 14 pancreatic cancer patients (mean age: 68 years old) during two visits. For six IVIM fit algorithms (IVIM-free, IVIM-adaptive, IVIM-Bayesian-log, IVIM-Bayesian-lin, IVIM-MLE and IVIM-fixed) and seven DWI models (IVIM, mono-exp, mono-exp-2, tri-exp, stretched-exp, Gaussian and kurtosis) the (a) goodness of fit (represented by adjusted R^2); (b) uniqueness of fit parameters (Spearman's rho); (c) precision (within-subject coefficient of variation, wCV); and (d) contrast between tumor and non-tumorous pancreatic tissue were assessed. A Wilcoxon signed-rank test ($\alpha = 0.05$) was performed between the best scoring parameter and remaining parameters (a, c, d), both for fit algorithms and DWI models.

Results: IVIM-Bayesian-lin and IVIM-fixed were overall the best IVIM fit algorithms. From the DWI models, Gaussian, kurtosis and monoexponential scored poorest. Excluding those, DWI models had high goodness of fit (a: $R^2 > 0.93$) and unique parameters (b). The perfusion fraction f_2 (tri-exp) had significantly higher tumor contrast (d) than all other parameters. Only the stretched-exp model had two precise (c: wCV < 15%) parameters, of which its perfusion-related parameter was significantly more precise than all other perfusion-related parameters.

Conclusion: IVIM-Bayesian-lin and IVIM-fixed are the best fit algorithms. The IVIM model scored relatively high in all tests, but for specific purposes, other models may be preferred.

Introduction

The intravoxel incoherent motion (IVIM) model for diffusion-weighted imaging (DWI) data obtained by magnetic resonance imaging (MRI) bears much promise as a tool to visualize tumors and monitor treatment response (e.g., in radiotherapy or chemotherapy) [1-3]. Contrary to the classical DWI model, in which signal attenuation is modeled monoexponentially as a function of diffusion-weighting (b-value), the IVIM model predicts a biexponential decay, probing both tissue diffusion and perfusion. Since the introduction of the IVIM model [4], the non-monoexponential behavior of DWI data in the pancreas was confirmed in multiple studies [1-3] and related to the interplay between diffusion and perfusion [5, 6]. Consequently, the IVIM model has been used to distinguish pancreatic cancer and surrounding tissue [3], characterized lesions in the pancreas [1, 2] and enabled treatment response monitoring in various other organs [7, 8].

Due to the limited precision of the IVIM parameters [9], multiple algorithms have been used to fit IVIM data. It was shown in simulations and healthy volunteer data that a Bayesian fit gives the most accurate and precise results in abdominal tissue [10]. However, thus far no analyses were done in pancreatic cancer patients. In these patients, fitting the data may be more challenging due to the limited size of the tumor compared to the entire organ, less perfusion [1-3], and echo planar imaging (EPI) artifacts that occur as a result of e.g. intratumoral fiducials [11] or biliary stents [12] close to the tumor.

The IVIM model is an approximation of the signal decay [13], and some of the assumptions underlying the IVIM model are not met in the pancreas [14]. Incoherent dephasing spins that result from different blood flow velocities within larger vessels can necessitate a third exponent in the IVIM model. Furthermore, rather than discrete (pseudo-) diffusion speeds, as used in the IVIM model, a smooth distribution of diffusion coefficients, reflecting the variation in cell sizes and vessel lengths within the voxel may be modeled. This was done in the stretched exponential [15] and Gaussian models [16].

Currently, there is no consensus on which fit algorithm for IVIM modeling and which DWI model should be used for different applications. We hypothesize that less common algorithms and fit models may outperform the most commonly used least squares fit of the mono-exponential and IVIM models for specific purposes. Hence, the objective of this exploratory study was to compare the performance of six IVIM model fitting algorithms and seven DWI models for pancreatic cancer imaging.

Materials and Methods

This prospective study (NCT01995240) was approved by the local ethics committee, and all patients gave written informed consent. Inclusion criteria were: locally advanced or metastatic pancreatic ductal adenocarcinoma, normal kidney function (eGFR > 60 mL/min/1.73m²) and no contraindication to undergoing MRI scanning. Sixteen consecutive patients fulfilling these criteria and willing to participate were included. Patients were scanned on a 3 T scanner (Philips Ingenia, Best, The Netherlands) between October 2014 and March 2016 at the Academic Medical Center in Amsterdam. One patient dropped out between scan sessions, and for one patient, the scans were stopped due to patient discomfort. Thus, data from fourteen patients were analyzed (eight females, mean age 67 years old, range 52–78, six males, mean age 70 years old, range 56–77). Data from nine of these patients were published previously [17].

Data acquisition

All patients were scanned three times during two separate sessions (average: 4.5 days apart, range: 1–8 days). To minimize bowel motion hyoscine bromide (Buscopan, Boehringer, Ingelheim, Germany; 20 mg IV) was administered directly before the first DWI acquisition in each session. The data from the acquisition without administration of hyoscine bromide were used for the intra-session analysis only.

For each patient, we acquired 2D multi-slice diffusion-weighted EPI data and contrast-enhanced (CE) T1-weighted multi-echo spoiled gradient echo (T1W GE) data with Dixon reconstruction (see Table 8.1 for detailed imaging parameters). The T1W GE data were acquired 35 seconds after Gadovist 1.0 (Bayer Healthcare, Leverkusen, Germany) administration (0.1 ml/kg; 5 ml/s, followed by 15 ml saline flush). Per b-value, data were acquired in isotropic distributed directions.

Post processing

All data analyses, fitting and statistical tests were performed in Matlab 2013a (MathWorks, Natick, U.S.A.), except for the IVIM-Bayesian-log fit, which was implemented in DTITools for Mathematica [18], Mathematica 10.4.1 (Wolfram Research, Champaign, U.S.A.).

All DWI images were denoised using a Rician adaptive non-local means filter [19] and registered in Elastix [20, 21] (Appendix I for details).

First, we tested the most common fitting algorithms for the IVIM model: IVIM-free, IVIM-adaptive [22], IVIM-Bayesian-log [23, 24], IVIM-Bayesian-lin [10], IVIM-MLE

[25] and IVIM-fixed (Appendix II for details). All IVIM model fit algorithms converted the IVIM signal fractions into volume fractions [6], using a $TE = 45$ ms and assuming a $TR = 5000$ ms (typical respiratory cycle), $T1 = 725$ ms and $T2 = 43$ ms for the pancreas and $T1 = 1932$ ms and $T2 = 275$ ms for blood [26, 27].

Then, we tested six additional common DWI models using the Levenberg-Marquardt least squares fit: the mono-exp, mono-exp-2, tri-exp, stretched-exp, Gaussian and kurtosis model (Appendix II for details). These were also compared to the two Levenberg-Marquardt least squares IVIM fits: IVIM-free and IVIM-fixed.

An abdominal radiologist (M.R.W.E., 9 years' experience) and an abdominal imaging researcher (R.K. 3 years' experience) drew regions of interest (ROIs) in consensus using 3D Slicer [28]. For each patient, two ROIs were created, one containing pancreatic tumor tissue and one containing non-tumorous pancreatic tissue. The ROIs were drawn on an ADC-map, generated from $b = 0$ mm²s and 600 mm²s, under the guidance of CE T1W GE images. The mean value of the voxel-wise fits within the ROIs was calculated.

Table 8.1. Sequence parameters.

	DWI	T1W GE
FOV (RL × AP) (mm ²)	432 × 108	400 × 353
Acquisition matrix	144 × 34	236 × 208
Slices	18	56
Slice thickness/gap (mm)	3.7/0.3	1.7/—
$TR^1/TE/\Delta TE$ (ms)	> 2200/45/—	4.7/1.15/1.0
FA (°)	90	10
BW (Hz/voxel)	59 (phase direction)	1602 (frequency)
Parallel imaging	1.3 (AP)	2/1.5 (RL/AP)
Partial Fourier	0.8	no
Respiratory compensation	Respiratory trigger (navigator)	1 breath-hold
Fat saturation	Gradient reversal during sliceDixon reconstruction selection + SPIR	
b-values (mm ² s) and directions/averages between brackets	0 (15), 10 (9), 20 (9), 30 (9), 40 (9), 50 (9), 75 (4), 100 (12), 150 (4), 250 (4), 400 (4) and 600 (16)	

¹ TR of the DWI acquisition was determined by the respiratory trigger interval, but it was at least 2200 ms. Abbreviations: DWI = diffusion-weighted imaging; FOV = Field of view; RL = right-left; AP = anterior-posterior; TR = repetition time; TE = echo time; ΔTE = increase in echo time; FA = flip angle; BW = bandwidth per voxel; SPIR = spectral presaturation with inversion recovery.

Comparison of methods

We quantitatively evaluated the performance of fitting algorithms and models considering the following four factors: the goodness of fit, the uniqueness of the fit parameters, the precision, and the fit parameters' contrast between tumor and non-tumorous pancreatic tissue.

Goodness of fit

For all models fitted with the Levenberg-Marquardt least squares fit, we took the mean adjusted R^2 from both ROIs as a measure for goodness of fit.

Uniqueness

For the IVIM fit algorithms and the multi-parametric DWI model, we used a Spearman's rank correlation test between the fit parameters to examine the unique nature of the fit parameters (significance level $\alpha = 0.05$). Fit parameter combinations with significant Spearman's rho indicate that measuring both parameters has limited added value. For this purpose, the fit parameters were averaged over the three acquisitions per patient.

Precision

From the repeated measures we calculated the inter- and intra-session within-subject coefficient of variation (wCV) of the tumor ROI as a measure of precision [29].

Tumor contrast

We tested whether parameter values from the tumor ROI (averaged over the three acquisitions per patient) were different from the values in non-tumorous pancreatic tissue (Wilcoxon signed-rank test; significance level $\alpha = 0.05$). Also, we calculated the contrast between tumor and non-tumorous pancreatic tissue. Per patient, the contrast was defined as the difference in parameter value of tumor and non-tumorous pancreatic tissue divided by the mean parameter value of both tissues multiplied by 100%. Finally, we plotted precision (wCV) as a function of contrast.

Statistical analyses

A Wilcoxon signed-rank test over the patient pairs was performed to determine which models had significantly lower adjusted R^2 compared to the model with highest adjusted R^2 , both for tumor and non-tumorous pancreatic tissue (significance level $\alpha = 0.05$). For wCV and contrast, tests were performed to compare the results between the different IVIM fit algorithms as well as between the different DWI models (all Levenberg-Marquardt least squares fits, including IVIM-free and IVIM-fixed). For the

IVIM fit algorithms, tests were performed per parameter (D , f , D^*). For the DWI models, parameters were split into two groups: diffusion-related parameters (D , ADC , ADC_{slow} , DDC , M_{ADC} , D_{app}) and perfusion-related parameters (f , f_1 , f_2 , D^* , α , σ_{ADC} , K_{app}). Tests were done between the best parameter of a group and the remaining parameters. To test whether parameters had higher wCV than the best parameter, a Wilcoxon signed-rank test was performed over the squared differences of the repeated measure (m_1 and m_2), divided by the squared mean (μ) of the population for that parameter: $(m_1 - m_2)^2 / \mu^2$ (significance level $\alpha = 0.05$). To test whether parameters had lower contrast than the best parameter, a Wilcoxon signed-rank test was performed over the contrast per patient (significance level $\alpha = 0.05$). Due to the exploratory nature of this research, no post hoc corrections were applied to the statistical tests presented.

Results

In two out of forty-two acquisitions it was not possible to delineate the tumor. Therefore, intra-session wCV s were determined using twelve patients. In one patient, no non-tumorous pancreatic tissue was present in the images. Therefore, tumor contrast was only based on thirteen patients. The mean mask sizes were 7.6 cm^3 ($= 210$ voxels, range 3.0 – 23.5 cm^3) for the tumor ROIs and 4.2 cm^3 ($= 118$ voxels, range 1.3 – 8.1 cm^3) for the non-tumorous pancreatic tissue ROIs.

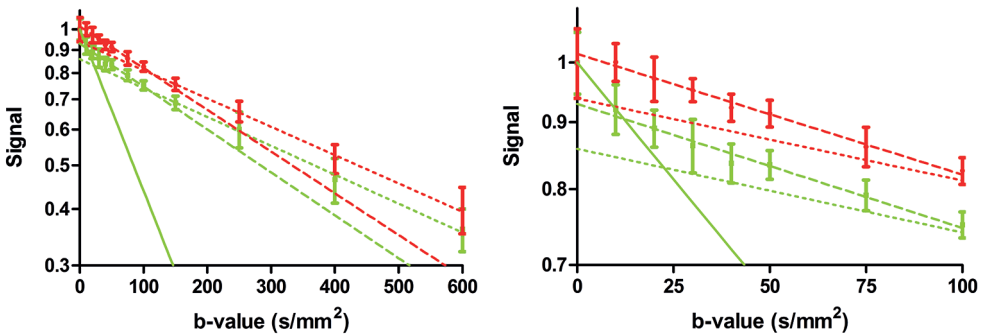


Figure 8.1. The normalized mean signal, averaged over all acquisitions, plotted as a function of the full range of b-values (**left**) and zoomed at low b-values (**right**) for signal from tumor (red) and pancreatic (green) ROIs. The lines indicate monoexponential fits over range $b = 150$ – $600 \text{ mm}^2\text{s}$ (dotted), $b = 10$ – $150 \text{ mm}^2\text{s}$ (striped) and, only for pancreatic tissue, $b = 0$ – $10 \text{ mm}^2\text{s}$ (solid). The figure shows that the classical IVIM-model does not fully describe the DWI data, in particular in pancreatic tissue.

Figure 8.1 highlights that the two exponents of the IVIM-model are insufficient to fully describe the DWI data, in particular in non-tumorous tissue, and underlines the importance of comparing several competing models. Parameter maps were generated for all fit algorithms and DWI models and shown in Fig. 8.2.

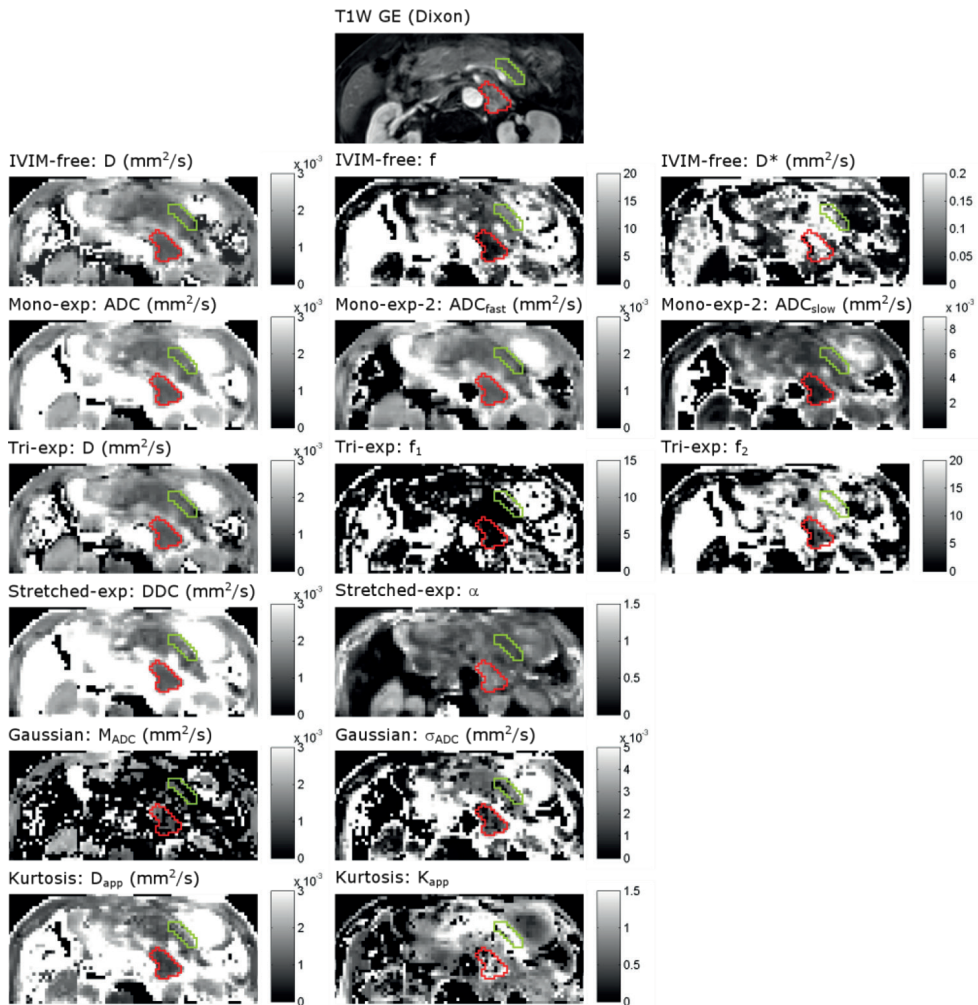


Figure 8.2. Axial parameter maps of the fit parameters of two IVIM model fit algorithms and seven DWI models in a 74 year-old female with pancreatic adenocarcinoma. ROIs containing diffusely growing pancreatic tumor infiltrating the adrenal gland (red) and pancreatic tissue in the pancreatic tail (green) are shown. The CE T1W GE is added as a reference.

Goodness of fit

The goodness of fit in tumor tissue was highest for the stretched-exp model (adjusted $R^2 = 0.96$; Fig. 8.3). However, all models had adjusted $R^2 > 0.93$ in tumor tissue. In non-tumorous pancreatic tissue the IVIM-free model had highest adjusted R^2 (0.94). Mono-exp, Gaussian and Kurtosis models had significantly lower adjusted R^2 in tumor as well as non-tumorous tissue and, in particular, they had adjusted $R^2 < 0.90$ in non-tumorous pancreatic tissue.

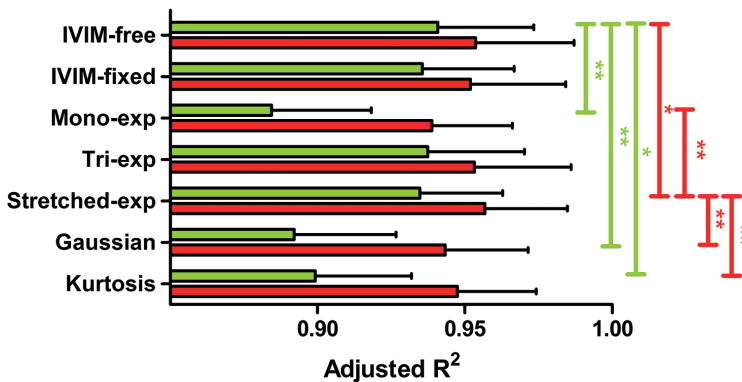


Figure 8.3. The mean adjusted R^2 in tumor (red) and pancreatic (green) tissue as a measure of goodness of fit. Bars indicate the standard deviation between patients. Lines indicate significant differences (*: $p < 0.05$, **: $p < 0.01$) between the model with highest adjusted R^2 in tumor (red) and pancreatic (green) tissue and the connected model. The p-values are given in Table 8.D from the Appendix IV.

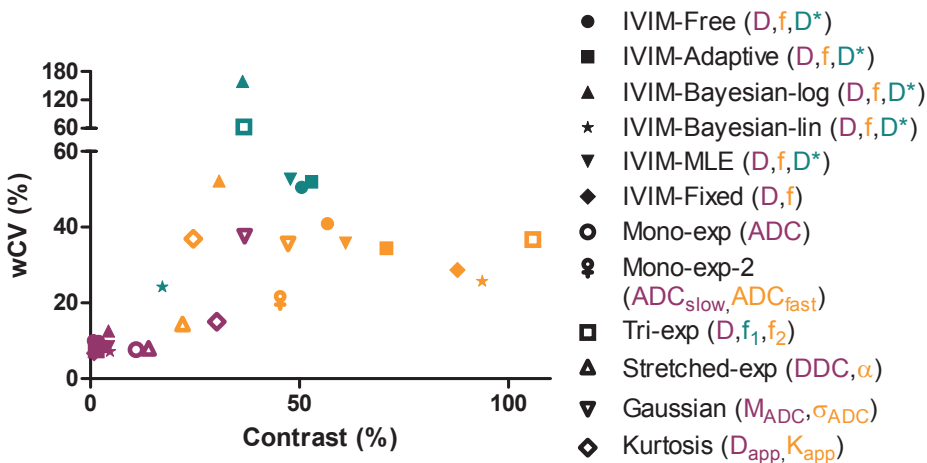


Figure 8.4. Plots of the wCV as a function of contrast between tumor and pancreatic tissue for the diffusion-related parameter (purple) and other fit parameters (orange, teal).

Uniqueness

Except for D & D^* and f & D^* of IVIM-adaptive, IVIM-Bayesian-log and IVIM-Bayesian-lin, there were no significant correlations between fit parameters (Table 8.2). The weakest correlations were found in the tri-exp model.

Precision

The inter-session wCVs were on average 30% larger than the intra-session wCVs (Table 8.3 and Table 8.C from Appendix III), indicating a larger test-retest variation when scans are repeated on separate days than in the same session. From the IVIM fit algorithms, IVIM-Bayesian-lin had most repeatable f and D^* , and the wCV for D was not significantly worse than the best wCV for D (IVIM-fixed). IVIM-fixed had most repeatable D , and its f was not significantly worse than f of the IVIM-Bayesian-lin. From the DWI models, D of IVIM-fixed had overall best wCV (6.7%) from the diffusion-related parameters and was significantly lower than D of IVIM-free, M_{ADC} and D_{app} . From the perfusion-related parameters α of the stretched-exp model was lowest (14.4%) and significantly better than all other perfusion-related parameters. The stretched-exp model was the only model that produced good wCV (< 15%) for both parameters.

Table 8.2. Uniqueness.

	Spearman's rho		
	D & f	D & D^*	f & D^*
IVIM-free	-0.09	0.45	0.24
IVIM-adaptive	0.12	0.57*	0.59*
IVIM-Bayesian-log	-0.48	0.68*	-0.81*
IVIM-Bayesian-lin	0.45	0.61*	0.67*
IVIM-MLE	0.10	0.37	0.47
IVIM-fixed	0.22		
	ADC_{slow} & ADC_{fast}		
Mono-exp-2	0.42		
	D & f_1	D & f_2	f_1 & f_2
Tri-exp	0.08	0.18	0.01
	DDC & α		
Stretched-exp	-0.36		
	M_{ADC} & σ_{ADC}		
Gaussian	-0.36		
	D_{app} & K_{app}		
Kurtosis	0.34		

* = significant correlation ($p < 0.05$)

Table 8.3. Precision of parameters and contrast.

	Inter-session wCV			Contrast (%)		
Fit algorithms	D	f	D^*	D	f	D^*
IVIM-free	10.0*	40.9	50.5**	0.7	56.7**	50.5
IVIM-adaptive	8.5*	34.4	51.9*	2.7	70.8**	52.9
IVIM-Bayesian-log	12.6	52.2 *	159.4*	4.3	30.8**	36.4
IVIM-Bayesian-lin	7.2	25.7	24.2	4.7	93.7	17.2**
IVIM-MLE	8.4*	35.8	52.7*	4.3	61.0**	47.8
IVIM-fixed	6.7	28.7		0.9	87.8**	
DWI models (least squares fits)	Diffusion-related	Perfusion-related		Diffusion-related	Perfusion-related	
	D	f	D^*	D	f	D^*
IVIM-free	10.0*	40.9**	50.5**	0.7	56.7**	50.5**
IVIM-fixed	6.7	28.7*		0.9	87.8**	
	ADC			ADC		
Mono-exp	7.7			10.9		
	ADC_{slow}	ADC_{fast}		ADC_{slow}	ADC_{fast}	
Mono-exp-2	8.83	20.61*		2.0**	45.4**	
	D	f_1	f_2	D	f_1	f_2
Tri-exp	7.6	63.2*	36.7*	1.4**	36.7**	105.8
	DDC	α		DDC	α	
Stretched-exp	8.0	14.4		13.9**	22.0**	
	M_{ADC}	σ_{ADC}		M_{ADC}	σ_{ADC}	
Gaussian	37.6**	35.6**		36.9	47.3**	
	D_{app}	K_{app}		D_{app}	K_{app}	
Kurtosis	15.0*	36.9**		30.2*	24.6**	

For both the diffusion-related parameter group as the perfusion-related parameter group, the parameter with lowest wCV is printed bold. Stars in indicate the parameters that were significantly (*: $p < 0.05$, **: $p < 0.01$) worse than the best scoring parameter of the five groups (D , f , and D^* for the algorithms; diffusion-related and perfusion-related for fit models). The p-values are reported in Table 8.E and 8.F from the Appendix IV. Abbreviations: wCV = within-subject coefficient of variation

Tumor contrast

Each fit method and algorithm, except for mono-exp, had at least one parameter that differed significantly between the tumor and non-tumorous pancreatic tissue (Table 8.4). From the fit algorithms, IVIM-Bayesian-lin was most promising, with a perfusion fraction that had significantly more contrast (contrast = 93.7%) than all

other algorithms (Table 8.3). From the DWI models, the greatest contrast (105.8%) was observed in f_2 of the tri-exp model, which was significantly larger than all other parameters (Table 8.3). Generally, parameters that showed a larger contrast between tumor and non-tumorous pancreatic tissue also had high wCVs (Fig. 8.4). However, the three parameters with the largest contrast (f : IVIM-Bayesian-lin; f : IVIM-fixed; f_2 : tri-exp) had a relatively low wCV compared to their high contrast.

Finally, the results of all five tests are visually summarized in Fig. 8.5.

Table 8.4(a). Parameter values in tumor tissue.

	Tumor tissue		
	D ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	f (%)	D^* ($10^{-3} \text{ mm}^2\text{s}^{-1}$)
IVIM-free	1.38 ± 0.14	$4.98 \pm 1.04^{**}$	$56.9 \pm 16.6^{**}$
IVIM-adaptive	1.41 ± 0.14	$3.94 \pm 1.04^{**}$	$58.0 \pm 17.4^{**}$
IVIM-Bayesian-log	1.36 ± 0.13	$7.56 \pm 2.09^*$	208.2 ± 106.9
IVIM-Bayesian-lin	1.41 ± 0.14	$2.56 \pm 0.81^{**}$	$83.5 \pm 10.2^{**}$
IVIM-MLE	1.43 ± 0.15	$4.42 \pm 0.99^{**}$	$61.1 \pm 17.1^{**}$
IVIM-fixed	1.50 ± 0.15	$2.60 \pm 0.75^{**}$	
	ADC ($10^{-3} \text{ mm}^2\text{s}^{-1}$)		
Mono-exp	1.60 ± 0.16		
	ADC_{slow} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	ADC_{fast} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	
Mono-exp-2	1.42 ± 0.14	$2.05 \pm 0.35^{**}$	
	D ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	f_1 (%)	f_2 (%)
Tri-exp	1.42 ± 0.14	$5.65 \pm 1.73^*$	$3.24 \pm 1.27^{**}$
	DDC ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	α	
Stretched-exp	$1.60 \pm 0.18^*$	$0.92 \pm 0.08^{**}$	
	M_{ADC} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	σ_{ADC} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	
Gaussian	0.87 ± 0.22	$1.68 \pm 0.33^{**}$	
	D_{app} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	K_{app}	
Kurtosis	$2.17 \pm 0.26^{**}$	0.91 ± 0.16	

Mean value \pm standard error ($1.96 \times SD/\sqrt{n}$; $n = 14$ for tumor, $n = 12$ for pancreatic). * = significantly ($p < 0.05$) different from pancreatic tissue (Wilcoxon signed-rank test on individual patients). ** = significantly ($p < 0.005$) different from pancreatic tissue. The p-values are reported in Table 8.E from the supplemental materials III.

Table 8.4(b). Parameter values in pancreatic tissue.

Pancreatic tissue			
	D ($10^{-3} \text{ mm}^2\text{s}^{-1}$)	f (%)	D^* ($10^{-3} \text{ mm}^2\text{s}^{-1}$)
IVIM-free	1.40 ± 0.12	8.22 ± 1.09	93.6 ± 21.3
IVIM-adaptive	1.45 ± 0.11	7.31 ± 1.04	98.1 ± 21.3
IVIM-Bayesian-log	1.43 ± 0.12	8.89 ± 1.49	246.3 ± 111.1
IVIM-Bayesian-lin	1.46 ± 0.12	6.38 ± 1.10	101.2 ± 13.4
IVIM-MLE	1.49 ± 0.13	7.57 ± 1.01	99.2 ± 21.4
IVIM-fixed	1.51 ± 0.12	6.12 ± 1.03	
ADC ($10^{-3} \text{ mm}^2\text{s}^{-1}$)			
Mono-exp	1.75 ± 0.19		
ADC_{slow} ($10^{-3} \text{ mm}^2\text{s}^{-1}$) ADC_{fast} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)			
Mono-exp-2	1.44 ± 0.13	3.05 ± 0.29	
D ($10^{-3} \text{ mm}^2\text{s}^{-1}$) f_1 (%) f_2 (%)			
Tri-exp	1.41 ± 0.12	7.28 ± 2.27	9.74 ± 2.47
DDC ($10^{-3} \text{ mm}^2\text{s}^{-1}$) α			
Stretched-exp	1.83 ± 0.23	0.74 ± 0.06	
M_{ADC} ($10^{-3} \text{ mm}^2\text{s}^{-1}$) σ_{ADC} ($10^{-3} \text{ mm}^2\text{s}^{-1}$)			
Gaussian	0.59 ± 0.14	2.51 ± 0.44	
D_{app} ($10^{-3} \text{ mm}^2\text{s}^{-1}$) K_{app}			
Kurtosis	2.83 ± 0.30	1.10 ± 0.12	

Mean value \pm standard error ($1.96 \times SD/\sqrt{n}$; $n = 14$ for tumor, $n = 12$ for pancreatic)

Discussion

We evaluated the performance of the six most common IVIM model fit algorithms and seven DWI models in patients with pancreatic cancer by assessing the goodness of fit, the uniqueness and precision of the fit parameters, and the contrast between tumor and non-tumorous pancreatic tissue. Among the IVIM fit algorithms, the IVIM-Bayesian-lin and IVIM-fixed perform best as they exhibit the highest precision and best contrast for the perfusion parameter f . Among the models, the IVIM model scored relatively high in all tests, but for specific purposes (e.g. diagnostics and treatment response monitoring) other models may be preferred.

IVIM fit algorithms

From the fit algorithms tested, IVIM-Bayesian-lin performed best considering precision and tumor contrast. IVIM-Bayesian-lin showed the highest precision for f and D^* compared to the other algorithms, and the precision on D was not significantly lower than for the algorithm with the highest precision (IVIM-fixed). Furthermore, the contrast between tumor and non-tumorous pancreatic tissue generated by f was significantly higher than in all other algorithms. These findings are in agreement with earlier published simulations and volunteer measurements [10]. However, IVIM-Bayesian-lin is not widely implemented. Also, the uniqueness of fit parameters was poor for all three combinations of parameters in both Bayesian approaches. Therefore, IVIM-fixed may be a good alternative. IVIM-fixed showed the highest precision of D , and the precision in f had second to best precision, which was not significantly worse than IVIM-Bayesian-lin. A disadvantage of IVIM-fixed is that no information on D^* is obtained. However, all algorithms had strong correlations between D^* and the other fit parameters, which implies a limited added value of D^* .

DWI models

The mono-exp, Gaussian, and kurtosis models overall scored poorly. This is explained by the fact that these models are unable to capture differences in the non-monoexponential behavior which separates tumour and non-tumorous pancreatic tissue. Therefore, these models had significantly lower adjusted R^2 compared to the DWI models with highest adjusted R^2 , both in tumor and non-tumorous pancreatic tissue. As the kurtosis model and Gaussian model were not developed to describe DWI data from b-values $< 150 \text{ mm}^2\text{s}$, it is interesting that they both had parameters that differed significantly between tumor and non-tumorous pancreatic tissue.

Overall, the IVIM model performed well, but for specific purposes, other models may be preferred. For diagnostics, when the goal is to distinguish between pancreatic tumors and non-tumorous pancreatic tissue, the model parameter with the highest contrast is preferred. We found that, in general, an increase in contrast came at the cost of a decrease in precision. However, for the parameters with highest contrast, the perfusion parameters of tri-exp (f_2), IVIM-fixed (f) and IVIM-Bayesian-lin (f), the increase in tumor contrast was to a much lesser extent accompanied by low precision. The f_2 parameter of tri-exp had significantly higher contrast than all other parameters, though the contrast of f of IVIM-fixed and IVIM-Bayesian-lin was still less than 20% lower. Therefore, these parameters are best at distinguishing between pancreatic tumors and non-tumorous pancreatic tissue.

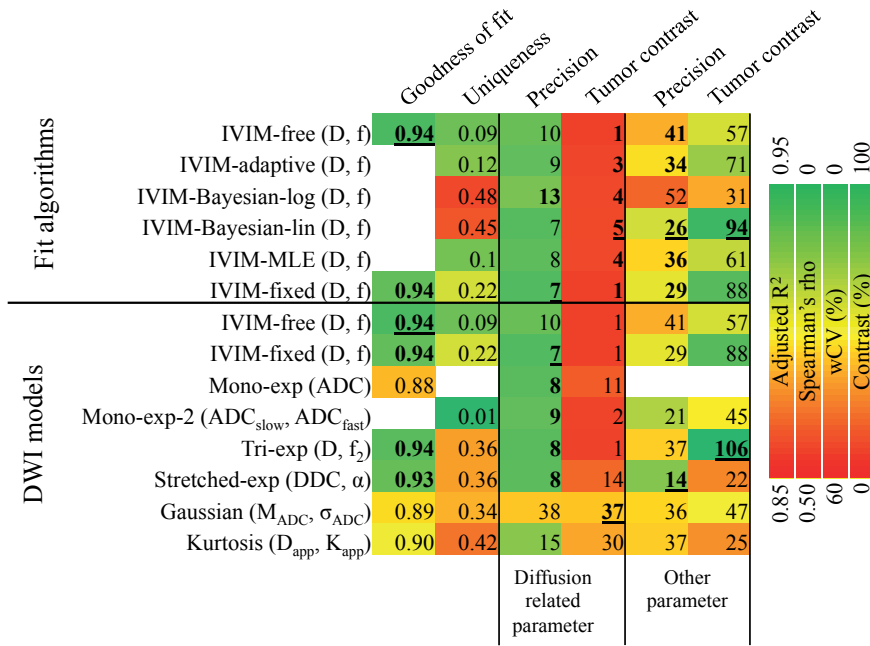


Figure 8.5. Summary of the results. The **left** column shows the goodness of fit represented by the adjusted R^2 from pancreatic tissue. The **second** column shows the uniqueness in tumor tissue, represented by the absolute value of Spearman's rho between the two parameters described in the remaining columns, illustrating the added value of the additional parameter. The **third** and **fourth** columns show the precision and tumor contrast of diffusion-related parameters D , ADC , ADC_{slow} , DDC , M_{ADC} and D_{app} . The **fifth** and **sixth** columns show the precision and tumor contrast of f , ADC_{fast} , f_2 , α , σ_{ADC} and K_{app} . Precision is represented by wCV. In the goodness of fit, precision and tumor contrast columns, the best scoring parameter is underlined, and parameters that were not statistically worse are bold.

For treatment evaluation and response monitoring, two aspects must be taken into account. First, the relevant model parameter needs to be measured with high precision and hence a low wCV. Second, there should be a change in the parameter of interest as a result of the treatment. The parameters with lowest wCV was D of the IVIM-fixed model, which is a diffusion related parameter. Tissue diffusion has been reported as a good biomarker for treatment response for e.g. chemotherapy of colorectal hepatic metastasis [30]. The wCV of D of IVIM-fixed was not significantly lower than wCV of many other diffusion related parameters: D (IVIM-Bayesian-lin, IVIM-Bayesian-log, tri-exp), ADC (mono-exp), ADC_{slow} (mono-exp-2) and DDC (stretched-exp). Consequently, several models can be used if diffusion is used as biomarker for treatment response. Perfusion-related parameters may be more sensitive to probe angiogenic changes as a result of therapy [7, 8]. In such a situation, f (IVIM) would be the most straightforward

to use, as it directly relates to a perfusing blood fraction. Still, one should be aware of the limited precision of most perfusion-related parameters. The latter is not the case for the parameter α of the stretched-exp model. This parameter is related to the non-monoexponential behavior of the data as induced by perfusion. Furthermore, it had significantly lower wCV than all perfusion-related parameters. Unfortunately, there is no straightforward connection between α and physical perfusion parameters. Nevertheless, changes in α should relate to angiogenic changes and thus α may be a good candidate to monitor angiogenic changes. Hence, the stretched-exp model could be most useful for treatment response monitoring with two precise parameters (DDC and α) probing both diffusion and perfusion effects. This is confirmed by the findings by Orton et al. [31].

Limitations

A limitation of this study is that the ROI delineations were based on CE T1W GE and ADC-maps from $b = 0$ and $600 \text{ mm}^2\text{s}$. Therefore, the ROI contained regions with low perfusion (CE T1W GE, ADC-map) and, potentially, diffusion restriction (ADC-map), which might have biased our perfusion measurements in tumors, compared to non-tumorous pancreatic tissues. However, so far, this is considered the best way to delineate pancreatic tumors.

We realize that with the inclusion of 14 patients, the accuracy of the estimation of the wCV is moderate (95% confidence interval is still 35% of the SD for $n = 14$). We believe that for comparing the different algorithms and models this number suffices, especially due to the paired nature of the comparison.

In this study, we chose to assess multiple fit algorithms only for the IVIM model and not for the other models. This focus is justified by the fact that currently IVIM is the most widely used model for DWI data next to the monoexponential model. In addition, for the other DWI models alternative fit algorithms have not been widely used previously. Potentially, the other models may also improve to some extent if alternative fits are used.

Conclusion

When IVIM modeling of DWI data from pancreatic cancer patients is desired, there are two preferred options: IVIM-Bayesian-lin and IVIM-fixed. Overall, the IVIM model scored well but, depending on the application, it may be useful to also consider the triexponential model (promising for tumor detection) or stretched exponent model (treatment response monitoring).

References

1. Klauf M, Mayer P, Bergmann F et al. Correlation of histological vessel characteristics and diffusion-weighted imaging intravoxel incoherent motion derived parameters in pancreatic ductal adenocarcinomas and pancreatic neuroendocrine tumors. *Invest. Radiol.* 2015; 50(11):792–797.
2. Lemke A, Laun FB, Klauf M et al. Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: Comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. *Invest. Radiol.* 2009; 44(12):769–775.
3. Re TJ, Lemke A, Klauf M et al. Enhancing pancreatic adenocarcinoma delineation in diffusion derived intravoxel incoherent motion f-maps through automatic vessel and duct segmentation. *Magn. Reson. Med.* 2011; 66(5):1327–1332.
4. Le Bihan D, Breton E, Lallemand D et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 1988; 168(2):497–505.
5. Heusch P, Wittsack HJ, Heusner T et al. Correlation of biexponential diffusion parameters with arterial spin-labeling perfusion MRI: Results in transplanted kidneys. *Invest. Radiol.* 2013; 48(3):140–144.
6. Lemke A, Laun FB, Simon D et al. An in vivo verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. *Magn. Reson. Med.* 2010; 64(6):1580–1585.
7. Gaeta M, Benedetto C, Minutoli F et al. Use of diffusion-weighted, intravoxel incoherent motion, and dynamic contrast-enhanced MR imaging in the assessment of response to radiotherapy of lytic bone metastases from breast cancer. *Acad. Radiol.* 2014; 21(10):1286–1293.
8. Koh DM. Science to practice: Can intravoxel incoherent motion diffusion-weighted MR imaging be used to assess tumor response to antivascular drugs? *Radiology* 2014; 272(2):307–308.
9. Gurney-Champion OJ, Froeling M, Klaassen R et al. Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas. *Invest. Radiol.* 2016; 51(4):211–220.
10. Barbieri S, Donati OF, Froehlich JM, Thoeny HC. Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs. *Magn. Reson. Med.* 2016; 75(5):2175–2184.
11. Gurney-Champion OJ, Lens E, van der Horst A et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med. Phys.* 2015; 42(5):2638–2647.
12. Gurney-Champion OJ, Bruins Slot T, Lens E et al. Quantitative assessment of biliary stent artifacts on MR images: Potential implications for target delineation in radiotherapy. *Med. Phys.* 2016; 43(10):5603–5615.
13. Kennan R, Gao J, Zhong J, Gore J. A general model of microcirculatory blood flow effects in gradient sensitized MRI. *Med. Phys.* 1994; 21(4):539.
14. Wetscherek A, Stieltjes B, Laun FB. Flow-compensated intravoxel incoherent motion diffusion imaging. *Magn. Reson. Med.* 2015; 74(2):410–419.
15. Bennett KM, Schmainda KM, Bennett RT et al. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* 2003; 50(4):727–734.
16. Yablonskiy DA, Bretthorst GL, Ackerman JJH. Statistical model for diffusion attenuated MR signal. *Magn. Reson. Med.* 2003; 50(4):664–669.
17. Gurney-Champion OJ, Klaassen R, Froeling M et al. Comparison of six different diffusion weighted MRI models in pancreatic cancer patients. *Proc. Intl. Soc. Mag. Reson. Med.* 24, Singapore: 2016:3434.
18. Froeling M, Nederveen AJ, Heijtel DFR et al. Diffusion-tensor MRI reveals the complex muscle architecture of the human forearm. *J. Magn. Reson. Imaging* 2012; 36(1):237–248.
19. Manjón J V, Coupé P, Martí-Bonmatí L et al. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J. Magn. Reson. Imaging* 2010; 31(1):192–203.

20. Klein S, Staring M, Murphy K et al. Elastix: A toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* 2010; 29(1):196–205.
21. Huizinga W, Poot DHJ, Guyader J-M et al. Non-rigid groupwise image registration for motion compensation in quantitative MRI. *Work. Biomed. Image Regist. LNCS vol. 8545*, 2014:184–193.
22. Wurnig MC, Donati OF, Ulbrich E et al. Systematic analysis of the intravoxel incoherent motion threshold separating perfusion and diffusion effects: Proposal of a standardized algorithm. *Magn. Reson. Med.* 2015; 74(5):1414–1422.
23. Orton MR, Collins DJ, Koh DM, Leach MO. Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven Bayesian modeling. *Magn. Reson. Med.* 2014; 71(1):411–420.
24. Dyvorne HA, Galea N, Nevers T et al. 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* 2013; 266(3):920–929.
25. Poot DHJ, Klein S. Detecting statistically significant differences in quantitative MRI experiments, applied to diffusion tensor imaging. *IEEE Trans. Med. Imaging* 2015; 34(5):1164–1176.
26. de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results. *Radiology* 2004; 230(3):652–659.
27. Stanisz GJ, Odrobina EE, Pun J et al. T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magn. Reson. Med.* 2005; 54(3):507–512.
28. Fedorov A, Beichel R, Kalpathy-Cramer J et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* 2012; 30(9):1323–1341.
29. Barnhart HX, Barboriak DP. Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl. Oncol.* 2009; 2(4):231–235.
30. Koh D-M, Scurr E, Collins D et al. Predicting Response of Colorectal Hepatic Metastasis: Value of Pretreatment Apparent Diffusion Coefficients. *Am. J. Roentgenol.* 2007; 188(4):1001–1008.
31. Orton MR, Messiou C, Collins D et al. Diffusion-weighted MR imaging of metastatic abdominal and pelvic tumours is sensitive to early changes induced by a VEGF inhibitor using alternative diffusion attenuation models. *Eur. Radiol.* 2016; 26(5):1412–1419.
32. Jensen JH, Helpert JA, Ramani A et al. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* 2005; 53(6):1432–1440.

Appendix I

DWI was acquired using respiratory triggering. All slices of the volume were acquired during one trigger to avoid inter-trigger mismatches between slices of one single volume (one b-value in one diffusion direction). This approach resulted in long acquisition periods each trigger (1.8 s). Patients were instructed to hold their breath during the typical noise produced by the EPI readout and to breathe freely during the navigator acquisition to minimize intra-trigger respiratory motion.

All DWI images were denoised using a Rician adaptive non-local means filter [19], with a search radius of three voxels and a patch radius of one voxel. The images were registered using a two-step approach in Elastix [20]. First, a reference image was created by averaging at least five manually selected acquisitions from the same respiratory position. To correct for bulk displacements between acquisitions, we performed a mutual information based rigid Euler transformation on each b-value to this reference image. The second step used a non-rigid b-spline registration based on mutual information to adjust for further deformations. During this step, we used two registration approaches: a single group-wise 4D [21] registration, or multiple 3D registrations for each b-value. We manually selected the registration approach that resulted in the most stable anatomy across the images acquired at different b-values and gradient directions.

Appendix II

Details on the IVIM fit algorithms are described in Table 8.A. Details of the fit models and their constraints are described in Table 8.B. For the tri-exp model, the pseudo-diffusion coefficients D^*_1 and D^*_2 were fixed to 0.014 and 0.093 mm²s⁻¹ (based on healthy pancreatic data in 16 volunteers, data not shown). Note that S_0 was fitted too, instead of fixed to value from the $b = 0$ mm²s acquisition. However, results of S_0 were not analyzed and S_0 further.

Per b-value we averaged the data to increase fitting speed. The mean data from b-values obtained in 4, 9, 12, 15 and 16 directions were weighted 1, 2, 3, 4 and 4 times during the fit, respectively. All fits were done voxel-wise.

Voxels with IVIM model parameter $f > 25\%$ contained mainly large vessels with instantaneous signal decay, which greatly influences f and D^* in the IVIM model [9]. For this reason, we removed voxels with $f > 25\%$ when determining mean f and D^* from the IVIM algorithms. In addition, when $f < 1\%$, there is too little perfusion signal to determine D^* . Therefore, only signal from voxels with $f > 1\%$ are considered when calculating mean D^* from ROIs.

We tested seven fit models in this study: the mono-exp, mono-exp-2, IVIM, tri-exp, stretched-exp, Gaussian and kurtosis model. The mono-exp model returned an apparent diffusion coefficient (ADC) value from a fit to all data, whereas the mono-exp-2 model returns an ADC_{slow} by fitting mono-exp to data from $b = 100 \text{ mm}^{-2}\text{s}$ and $b = 600 \text{ mm}^{-2}\text{s}$ and an ADC_{fast} by fitting mono-exp to data $b = 0 \text{ mm}^{-2}\text{s}$ and $b = 100 \text{ mm}^{-2}\text{s}$. Mono-exp-2 probes IVIM-like features as ADC_{slow} represents diffusion and ADC_{fast} is also sensitive to perfusion but does not require as long a measurement time as typically seen for IVIM modeling. The tri-exp model described data from three different compartments with different (pseudo-) diffusion speeds (D , D^*_1 and D^*_2). D^*_1 relates to capillary perfusion, and D^*_2 relates to instantaneous dephasing of spins. The stretched-exp model [15] as well as the Gaussian model [16], which were also fitted, describe a continuous distribution of diffusivities rather than discrete diffusion values within a voxel. The Gaussian model does not imply the assumption that molecules diffuse Gaussian, but assumes that within a voxel, there is a single Gaussian distribution of ADC-values. Finally, the kurtosis model allows to incorporate non-Gaussian behavior of restricted diffusing molecules [32]. All models had parameters constrained to an appropriate domain (Table 8.B).

Table 8.A. Fit algorithms.

Name	Fit
IVIM-free	Levenberg-Marquardt least squares
IVIM-adaptive [22]	Adaptive threshold segmented fit
IVIM-Bayesian-log* [23, 24]	Data-driven Bayesian fit for which the prior is a fitted Gaussian in log-space to confine parameters to relevant values
IVIM-Bayesian-lin* [10]	Data-driven Bayesian fit using boxcar functions with support over pre-defined ranges as weakly informative priors.
IVIM-MLE [25]	Maximum likelihood estimator approach that assumed Rician noise
IVIM-fixed	Levenberg-Marquardt least squares, except that D^* was fixed to $70 \times 10^{-3} \text{ mm}^2\text{s}^{-1}$, which resulted in more stable fits in healthy volunteers [9] (value based on volunteer data).

* Both Bayesian fits required a region of interest (ROI) containing data used for the data-driven fit. To determine the prior, a ROI larger than typical tumor size was required and hence ROIs were selected containing the tumor, pancreas, and surrounding tissue

Table 8.B. DWI models.

	Formula	Parameters	Constraints
Mono-exp	$\frac{S(b)}{S_0} = \text{Exp}[-ADC \cdot b]$	ADC = apparent diffusion coefficient	—
IVIM ^a [4]	$\frac{S(b)}{S_0} = (1-f) \cdot \text{Exp}[-D \cdot b] + f \cdot \text{Exp}[-(D+D^*) \cdot b]$	D = diffusion coefficient D^* = pseudo diffusion coefficient f = perfusion fraction	0.5×10^{-3} – 6×10^{-3} mm ² s ^b 6×10^{-3} – 200×10^{-3} mm ² s 0.1–99%
Tri-exp	$\frac{S(b)}{S_0} = (1-f_1-f_2) \cdot \text{Exp}[-D \cdot b] + f_1 \cdot \text{Exp}[-D_1^* \cdot b] + f_2 \cdot \text{Exp}[-D_2^* \cdot b]$	D = diffusion coefficient D_1^* = pseudo diffusion coefficient 1 D_2^* = pseudo diffusion coefficient 2 f_1 = tissue fraction with D_1^* f_2 = tissue fraction with D_2^*	0 – 10×10^{-3} mm ² s 14×10^{-3} mm ² s 93×10^{-3} mm ² s 0.1–70% 0.1–70%
Stretched-exp [15]	$\frac{S(b)}{S_0} = \text{Exp}[-(b \cdot DDC)^\alpha]$	DDC = distributed diffusion coefficient α = stretching coefficient	0 – 6×10^{-3} mm ² s 0 – 2
Gaussian [16]	$\frac{S(b)}{S_0} = \frac{1 + \phi\left(\frac{M_{ADC}}{\sigma_{ADC} \cdot \sqrt{2}} - \frac{b \cdot \sigma_{ADC}}{\sqrt{2}}\right)}{1 + \phi\left(\frac{M_{ADC}}{\sigma_{ADC} \cdot \sqrt{2}}\right)} \cdot \text{Exp}\left[-b \cdot M_{ADC} + \frac{1}{2} \sigma_{ADC}^2 b^2\right]$	M_{ADC} = apparent diffusion coefficient distribution maxima σ_{ADC} = ADC distribution width	0.01 – 10×10^{-3} mm ² s 0.1 – 100×10^{-3} mm ² s
Kurtosis [32]	$\frac{S(b)}{S_0} = \text{Exp}\left[-b \cdot D_{app} + \frac{1}{6} b^2 D_{app}^2 K_{app}\right]$	D_{app} = apparent diffusion coefficient K_{app} = apparent diffusion kurtosis	0.5×10^{-3} – 14×10^{-3} mm ² s 0 – 10

$S(b)$ is the signal from an acquisition with b-value b , S_0 is the signal with no diffusion-weighting and ϕ is the Gaussian error function. ^a This is the simple IVIM formula; we implemented the TR , TE , $T1$ and $T2$ corrected version [6]. ^b D had no constraints in the IVIM-adaptive approach; IVIM-Bayesian-log had the following constraints: $D > 0$ mm²s, $D^* > 0$ mm²s and $0\% < f < 100\%$.

Appendix III

Table 8.C. Intra-session wCV.

	Intra-session wCV		
Fit algorithms	D	f	D^*
IVIM-free	6.3	32.9	32.7
IVIM-adaptive	5.3	33.9	39.3
IVIM-Bayesian-log	9.8	42.7	112.0
IVIM-Bayesian-lin	5.8	27.7	26.3
IVIM-MLE	5.0	38.7	41.8
IVIM-fixed	5.0	23.8	
	ADC		
Mono-exp	4.5		
	ADC_{slow}	ADC_{fast}	
Mono-exp-2	6.32	18.04	
	D	f_1	f_2
Tri-exp	4.8	47.7	37.0
	DDC	α	
Stretched-exp	3.8	6.9	
	M_{ADC}	σ_{ADC}	
Gaussian	27.4	23.9	
	D_{app}	K_{app}	
Kurtosis	7.1	24.2	

Abbreviations: wCV = between-subject coefficient of variation

Appendix IV

Table 8.D. The p-values (Wilcoxon signed-rank) for the adjusted R^2 compared to the best.

	Adjusted R^2	
	Tumor	Pancreatic
IVIM-free	0.017	Best
IVIM-fixed	0.946	0.173
Mono-exp	< 0.001	0.005
Tri-exp	0.168	0.091
Stretched-exp	Best	0.153
Gaussian	< 0.001	< 0.001
Kurtosis	< 0.001	0.042

Table 8.E. The p-values for the inter-session wCV.

	Inter-session wCV		
Fit algorithms	D	f	D^*
IVIM-free	0.030	0.463	0.007
IVIM-adaptive	0.020	0.104	0.017
IVIM-Bayesian-log	0.104	0.035	0.035
IVIM-Bayesian-lin	0.153	Best	Best
IVIM-MLE	0.042	0.268	0.011
IVIM-fixed	Best	0.761	
DWI models (least squares fit)	Diffusion-related	Perfusion-related	
	D	f	D^*
IVIM-free	0.030	0.007	0.007
IVIM-fixed	Best	0.035	
	ADC		
Mono-exp	0.325		
	ADC_{slow}	ADC_{fast}	
Mono-exp-2	0.135	0.042	
	D	f_1	f_2
Tri-exp	0.241	0.001	0.049
	DDC	α	
Stretched-exp	0.194	Best	
	M_{ADC}	σ_{ADC}	
Gaussian	< 0.001	< 0.001	
	D_{app}	K_{app}	
Kurtosis	0.013	< 0.001	

Data are split into 5 groups (D , f and D^* for the algorithms; diffusion-related and perfusion-related for fit models). The p-values of the Wilcoxon signed-rank test between the parameter with lowest wCV (Best) of each group and the rest of that group are noted. Abbreviations: wCV = between-subject coefficient of variation

Table 8.F. The p-values for the contrast parameters.

	Tumor and pancreatic			Contrast		
Fit algorithms	D	f	D^*	D	f	D^*
IVIM-free	0.424	0.005	0.021	0.305	< 0.001	0.787
IVIM-adaptive	0.233	0.016	0.003	0.685	< 0.001	Best
IVIM-Bayesian-log	0.110	0.021	0.007	0.080	< 0.001	0.414
IVIM-Bayesian-lin	0.151	0.003	0.016	Best	Best	< 0.001
IVIM-MLE	0.339	0.016	0.009	0.216	< 0.001	0.191
IVIM-fixed	0.380	0.007		0.787	0.008	
ADC models (least square fit)				Diffusion-related	Perfusion-related	
				D	f	D^*
IVIM-MLE				< 0.001	< 0.001	0.002
IVIM-fixed				< 0.001	< 0.001	
	ADC			ADC		
Mono-exp	0.110			< 0.001		
	ADC_{slow}	ADC_{fast}		ADC_{slow}	ADC_{fast}	
Mono-exp-2	0.266	< 0.001		< 0.001	< 0.001	
	D	f_1	f_2	D	f_1	f_2
Tri-exp	0.569	0.016	0.009	< 0.001	< 0.001	Best
	DDC	α		DDC	α	
Stretched-exp	0.052	0.021		< 0.001	< 0.001	
	M_{ADC}	σ_{ADC}		M_{ADC}	σ_{ADC}	
Gaussian	0.176	0.007		Best	< 0.001	
	D_{app}	K_{app}		D_{app}	K_{app}	
Kurtosis	< 0.001	0.301		0.017	< 0.001	

Left: p-value for Wilcoxon signed-rank test between parameter values in tumor and pancreatic tissue.

Right: data are split into 5 groups (D , f , and D^* for the algorithms; diffusion-related and perfusion-related for fit models). The p-value for Wilcoxon signed-rank test comparing contrasts from the parameter with the highest contrast (Best), compared to the remaining parameters per group are noted.



CHAPTER 9

General Discussion

Oliver J Gurney-Champion

In radiotherapy, treatment plans are based on tumor delineations. As these treatment plans are used for all radiotherapy sessions, the delineations should be as accurate as possible. Currently, such delineations are often based on computed tomography (CT) images. In **chapter 2** it was shown that there is a large interobserver variability in CT-based delineations of pancreatic cancer patients. In **chapter 3** the interobserver variability decreased when MRI images were made available during tumor delineation on CT. The interobserver variability may further improve when CT and MRI images are registered, which was not done in **chapter 3**.

In order to register MRI to CT, it is desirable to have intratumoral fiducial markers that are visible on CT and MRI. The visibility of golden fiducial markers and their potential to cause artifacts was quantified in **chapter 4**, using a sequence-independent approach. It was shown that at least one of the commercially available marker types was visible *in vivo*, in a pancreatic cancer patient. However, markers that were visible also caused artifacts, especially on diffusion-weighted imaging (DWI). The artifacts caused by such markers were smaller than artifacts caused by typical biliary stents. As 70% of pancreatic cancer patients receive such biliary stents, the artifacts caused by several metal stents and a plastic stent were quantified in **chapter 5**.

In **chapter 6** an alternative sequence for the T2-weighted turbo spin echo (T2W TSE) was optimized for abdominal use: the alternating repetition time balanced steady state free precession (ATR-SSFP) sequence. Contrary to T2W-TSE, this sequence allowed for isotropic high resolution ($1.4 \times 1.4 \times 1.4 \text{ mm}^3$) imaging within one breath-hold.

Despite encouraging results of radiotherapy treatment (neo-)adjuvant to surgery in pancreatic cancer patients, not all tumors respond to radiotherapy. As radiotherapy is a toxic treatment, it is desirable to only treat patients with tumors that are sensitive to radiotherapy. To avoid unnecessary irradiation in patients with non-responding tumors, treatment response monitoring or, ultimately, treatment response prediction is desirable. The intravoxel incoherent motion (IVIM) model for DWI potentially enables this. However, DWI acquisitions for IVIM data are often long and little is known about the optimal acquisition scheme and fitting. Therefore, in **chapter 7** the acquisition scheme for DWI that renders the minimum acquisition time necessary for accurate and precise IVIM modeling was determined. In **chapter 8** the best fit models and fitting algorithms for DWI data from pancreatic cancer patients were determined.

In **chapter 9** (this chapter), the work is contextualized to a larger perspective. First, the research in addition to the work in **chapter 2** and **chapter 3** that is required to determine appropriate treatment margins for pancreatic cancer patients is discussed. Then, the work from **chapter 4** and **chapter 5** is dealt with in the context of how it could be implemented to better understand artifacts. Following this, the

importance of having visible markers and the artifacts caused by metal markers is discussed. Also, the future of DWI in tumor imaging is discussed in light of the results from **chapter 7** and **chapter 8**. Furthermore, several 4DMRI techniques that allow monitoring of tumor positions throughout the respiratory cycle are considered. These techniques could also be used to generate higher resolution MRI as they are not limited to a single breath-hold. Finally, MRI-guided radiotherapy is discussed.

Treatment margins

In this thesis the interobserver variation, expressed as standard deviation (SD), of pancreatic cancer target volume delineation on CT (**chapter 2**) and CT+MRI (**chapter 3**) was assessed. These SDs are often used to determine treatment margins [1, 2]. However, there are several reasons why the overall observer variation mentioned in **chapter 2** and **chapter 3** are not ideal for determining treatment margins. For instance, common margin formulas are based on a normal distribution of errors. In pancreatic cancer, there were many locations at which delineation errors were not normally distributed. For example, some observers included a stent or a lymph node in the gross tumor volume (GTV), and others did not. In these cases, the variation was not Gaussian (i.e. either included or not included) and a treatment margin will not suffice to incorporate such errors. These errors were the result of an incomplete delineation protocol and poor observer compliance [3-5]. Therefore, clearer delineations guidelines should be introduced to achieve a more precise delineation.

In **chapter 3** it was shown that CT+MRI-based delineations often result in smaller overall SDs than CT-only based delineations. Similar findings were reported for several other organs too [6-9]. These findings suggest that more precise pancreatic cancer delineations are achieved when MRI images are available during delineation. As a result, smaller planning target volume (PTV) treatment margins may be used when delineations are based on CT+MRI than when they are based on CT.

However, in **chapter 3** as well as in other studies in pancreatic [10, 11] and other tumors [6-8], delineations based (partially) on MRI were smaller than delineations based on CT. Therefore, it is important to know how well the delineation represents the tumor size. So far, a study with one observer showed that the tumor size was underestimated by 4 mm on contrast-enhanced (CE) MRI when compared with pathology [12]. Such underestimations are a result of microscopic extensions of the tumor and are commonly incorporated in the clinical target volume (CTV) margin [13]. To determine appropriate CTV-margins, more elaborate multi-observer studies are required that compare delineations with pathology. Interestingly, the size of the

delineated tumor is different when delineated on images from different MRI sequences [10]. This suggests that the required CTV margins may be sequence dependent.

Artifacts

Metal implants can cause susceptibility artifacts in MRI images. Such artifacts can obscure the underlying pathology or deform the anatomy in the image. In many applications, including radiotherapy, such artifacts are undesired. Therefore, many studies have investigated artifacts caused by patient anatomy [14, 15], fiducial markers [16], stents [17-19] and other implants [20-22]. These studies described the artifacts for one specific sequence. Consequently, their results are often limited to the sequence presented. More generalizable results can be obtained if such artifacts are studied using a sequence independent approach. Such an approach was implemented in **chapter 4** and **chapter 5**. In those chapters, the usefulness of the approach was already illustrated for markers and stents. In a separate study [23], we also verified the applicability of the sequence independent approach for a brachytherapy applicator for cervical cancer.

In this sequence independent approach, the parameters underlying the artifacts, ΔB_0 and T_2^* , are deduced from MRI measurements. Alternatively, these quantities can be simulated if the shape and materials of the implants are known [24-27]. Such an approach is very useful for developing new implants.

Once the ΔB_0 - and T_2^* -maps around the implants are obtained either from simulations or measurements, they can be used to simulate artifacts [24-28]. Using this method, one can discriminate the extent to which specific anatomical features are contaminated by artifacts in the real image. Such simulated artifacts are particularly useful when based on ΔB_0 - and T_2^* -maps acquired *in vivo*, in the patient being studied [28]. The data required for such ΔB_0 -maps and T_2^* -maps can be obtained within a 20-second breath-hold. This *in vivo* approach was explored in the supplementary materials from **chapter 5**.

Finally, there are many ways to reduce artifacts. When a ΔB_0 -map is acquired, the artifact can be reverted by applying a transformation which is the inverse of the expected deformation (from the ΔB_0 -map) [29-31]. Also matching to a reference image can undo deformations [32]. However, these methods cannot solve signal summation (i.e. signal from multiple voxels shifting to the same voxel). Alternatively, by acquiring multiple images with different read-out directions it becomes possible to undo deformations and signal summation in some cases [33, 34]. Also, the sensitivity of the sequence to such artifacts can be reduced during acquisition [35]. All these

methods, however, have their limitations and often come at the cost of additional scan time. This is especially challenging when scanning the abdomen, where scan duration is restricted by breath-holding.

Fiducial markers

In radiotherapy, matching MRI, CT and CBCT is important to ensure accurate treatment. Matching images requires visualization of several landmarks in the patient. In some cases, these landmarks can be anatomical, such as bony anatomy for matches between CT and CBCT [36] or major vessels in matches between MRI and ultrasound [37]. However, the location of several organs (e.g. pancreas [38] and prostate [39, 40]) with respect to the bony anatomy, varies over time. Furthermore, due to the poor soft tissue contrast on CT, and in particular on CBCT, matching other anatomical features, such as major vessels, remains challenging.

Therefore, matching CT/CBCT to different MRI images of such organs requires intratumoral fiducial markers to ensure alignment of the tumor [38, 39, 41]. From **chapter 4** it is clear that for golden fiducial markers the visibility is related to the marker's potential to cause artifacts. Therefore, all gold markers that were visible on MRI in earlier studies [42-45] will probably also cause artifacts on artifact sensitive sequences such as sequences with EPI read-out [16]. Furthermore, it is likely that all markers in which visibility is based on local $T2^*$ decay, will also cause signal shifting artifacts. As a result, when marker visibility is desired without artifacts, an alternative approach is required.

Alternative markers that are visible on MRI and CT have been proposed, such as hydrogel [46, 47] or a mixture of iodinate contrast agent and water [48]. The hydrogel was tested *in vivo* in other organs [46, 47]. The iodinate mixture was tested *in vitro*. Both fit through a 22 gauge needle and hence might be implantable in pancreatic cancer patients [49]. However, many factors still need to be investigated, such as the stability of those gels *in vivo*.

Diffusion-weighted imaging

DWI and other functional imaging, such as dynamic contrast-enhanced (DCE) and $T2^*$ acquisitions, may help in predicting treatment outcome or enabling treatment monitoring [50, 51]. Correlating MRI parameters, including DWI model parameters, to treatment outcome of neoadjuvant chemoradiotherapy and resection of pancreatic

cancer is being investigated in the MIPA study at the AMC (NCT01989000). Potentially, DWI can be used to determine which patients benefit from neoadjuvant chemoradiotherapy in the future.

Furthermore, DWI and other functional imaging may be used to detect, characterize and map different tumor regions. More aggressive or radiation resistant parts of the tumor could then be selected to receive a higher radiation dose. This approach is called dose painting [52-54]. Dose painting in the pancreas, however, would require sophisticated motion management techniques to prevent the high dose from blurring. Furthermore, a better observer agreement on the tumor location (**chapter 2** and **chapter 3**) should be achieved before dividing the tumor into smaller subsections.

DWI is often obtained at 1.5 T and 3 T [55]. All DWI data presented in this thesis were obtained at 3 T. DWI has low SNR and therefore higher field strengths may improve DWI. However, the echo planar imaging (EPI) read-out associated with DWI is very sensitive to susceptibility artifacts, which are accentuated at higher field strength. Alternatively, read-outs that are less sensitive to artifacts can be used, such as diffusion prepared TSE or bSSFP [56-58]. When used at higher field strengths, such sequences benefit from added SNR without the drawback of artifacts. We illustrated the feasibility of a diffusion prepared stimulated TSE sequence in the prostate [59]. Alternatively, a diffusion-weighted steady state sequence could be used [60, 61]. The challenge that comes with the application of the sequences mentioned above for abdominal use is their strong sensitivity to motion combined with acquisitions much longer than a breath-hold (typically > 5 minutes).

Competing models

Before performing treatment monitoring or dose painting, one needs to select a DWI model. Depending on the purpose, several factors should be considered when selecting a model. E.g., when depicting tumors is the goal, the model with largest tumor contrast is desired, whereas when treatment response monitoring is the goal, a model with good test-retest stability is desired. Currently, the monoexponential decay model and IVIM model are most popular for abdominal DWI [55, 62-65]. There are many competing DWI models and therefore, in **chapter 8** their performances were compared.

Classically, DWI data is modelled as a monoexponential decay as a function of diffusion weighting, in which the exponent is the apparent diffusion coefficient. The IVIM model describes the signal decay as a biexponential decay as a function of diffusion weighting. One exponent is assumed to relate to diffusion in tissue and

one to capillary blood [66]. The non-monoexponential behavior of DWI data was illustrated in multiple studies and related to perfusion [67, 68]. In the triexponential model, blood from large vessels is also modeled. Another alternative model is the stretched exponent model [69]. This model is a data-driven model with no thorough physical model behind it, and hence the parameters have no straight forward physical meaning.

In **chapter 8** it was shown that for DWI of pancreatic cancer, the IVIM, stretched exponent and triexponential models performed best of the seven models tested. In other research [70, 71] we illustrated the necessity of a triexponential model to accurately relate model parameters to kidney function. Also, it was shown for multiple organs that the stretched exponent model was more precise [72, 73], more sensitive to treatment response [74] or better at tumor classification [75] than the monoexponential and IVIM models. The fact that different models describe the data better in different organs and for different applications illustrates that the underlying mechanism of diffusion is not yet fully understood.

We believe that part of the discrepancy between the IVIM model and the data is due to incoherent bulk motion (e.g. blood in large vessels flowing at different speeds). Such bulk motion would cause additional signal decay, which in the exponential model approach necessitates a third exponent. Signal decay from incoherent bulk motion can be eliminated by using motion compensated diffusion gradients [76]. This way, the data decay should follow a more biexponential shape. However, whether this approach would be clinically beneficial is doubtful. The approach would eliminate part of the signal related to the ultra-fast signal decay in the triexponential model. We found that the parameter related to this ultra-fast signal decay showed the largest contrast between pancreatic cancer and healthy pancreatic tissue (**chapter 8**) in one study and correlated with kidney perfusion in another study [71].

Faster acquisitions

Acquiring data for IVIM modelling costs time, as data from several diffusion-weighted are required. Therefore, in **chapter 7**, we investigated the shortest acquisition scheme available to obtain data that allowed for accurate and precise IVIM model fitting. In this chapter, data from healthy volunteers was used. To see if this also holds true in patient data, we simulated the optimal accelerated acquisition scheme from **chapter 7** in our patient data from **chapter 8** by removing the redundant data. This accelerated acquisition had a within subject coefficient of variation that was less than 20% higher than the full acquisition for all parameters from the IVIM model. As an increase of the within subject coefficient of variation of less than 20% was the criteria

for the accelerated acquisition in **chapter 7** we can conclude that the scheme holds in patients.

4DMRI and high resolution

In radiotherapy, it is desirable to know the location of the tumor throughout the respiratory cycle. Therefore, for abdominal tumors, a 4DCT is often obtained on which the tumor motion can be tracked and included in the GTV. For example, in **chapter 2** and **chapter 3** radiation oncologists delineated the tumor on 4DCTs. A 4DCT consists of several sets (typically 10) of CT images, each representing a part of the respiratory cycle. Tumor visibility is often poor on such CTs. As MRI has stronger tumor contrast, a similar MRI (4DMRI) could help improve tracking tumor motion over the respiratory cycle. There are several approaches to generate such a 4DMRI.

A straightforward implementation is by obtaining fast 2D slice images, for example by a single shot TSE or EPI acquisition, during free breathing. If each slice acquisition is interleaved with a navigator acquisition tracking the liver dome, the respiratory cycle can be monitored. The 2D slice images can be sorted retrospectively according to respiratory phase. We showed the feasibility of this approach for pancreatic imaging [77]. This approach was relatively simple and robust; however, several limitations were associated with it. The acquisition was limited to 2D acquisitions, which resulted in a small number of thick slices. This is undesirable when delineating tumors in radiotherapy. Furthermore, the acquisitions were long (approximately 7 minutes). Also, the reconstruction often missed data for certain slice-respiratory phase combinations. In a variation to this approach, acquisitions are triggered at certain

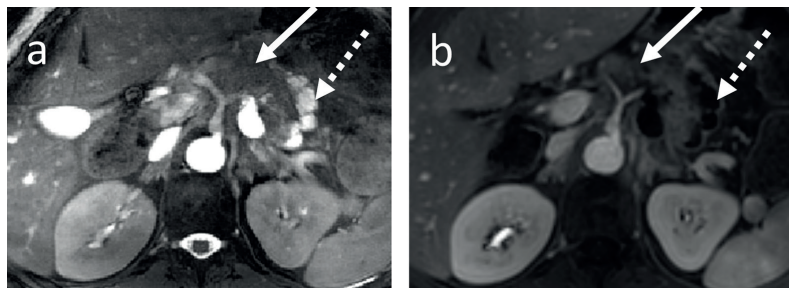


Figure 9.1. A frame of a 4DMRI from a pancreatic cancer patient obtained with ATR-bSSFP (**a**). MRI was obtained using the tiny golden-angle radial stack of stars ATR-bSSFP acquisition. A 3D CE T1-weighted gradient echo image is added as reference (**b**). Solid arrow indicates solid tumor region whereas the dotted arrow indicates cysts/necrotic region.

respiratory phases to ensure data completeness [78, 79]. However, the prospective nature of that approach requires determining the amplitude of the respiratory cycle at an early stage during acquisition. This makes the prospective method more sensitive to irregular breathing. Furthermore, these acquisitions were reported to last up to 18 minutes.

Alternatively, a 4DMRI acquisition can be achieved with a tiny golden angle [80, 81] radial stack of stars gradient echo acquisition, obtained during free breathing. The respiratory signal can then be obtained from the center of k-space for each stack of k-lines at a given angle. The respiratory signal can be used to sort the k-space data in different respiratory phases. This allows reconstructing high-resolution 3D images (instead of multiple 2D slices) for every respiratory phase [82]. Furthermore, by using a compressed sensing reconstruction [83] the acquisition can be accelerated to obtain 4DMRI to within a minute [84]. This approach can be applied to most gradient echo sequences and therefore allows for a variety of contrasts. As an extension of the work discussed in **chapter 6**, we applied this 4DMRI technique to the ATR-SSFP sequence to obtain high-resolution 4D ATR-SSFP images (Fig. 9.1) [85]. Furthermore, we are working on a similar approach to generate 4D $T2^*$ -maps to quantify $T2^*$ values in abdominal lymph nodes. In such an application, acquiring images during breath-holding does not allow for sufficient resolution to discern the lymph nodes. These k-space sorting techniques can be extended further to generate MRI from all respiratory and cardiac phases [86].

Future technologies

MRI-guided radiotherapy

The recent development of several radiotherapy treatment systems with integrated MRI [87-90] enabled MRI-guided radiotherapy (MR-IGRT). Currently, the only commercially available and clinically released MR-IGRT system is the MRIdian system (ViewRay, Oakwood Village, USA) [87]. This system uses a 0.35 T MRI combined with three Cobalt-60 irradiation units perpendicular to the main magnetic field. A major point of criticism on this system was the use of Cobalt sources instead of a linear accelerator. However, a linear accelerator (linac) module upgrade of this system recently received CE mark approval.

There are three alternative systems that all use a 6 MV linac but are not yet clinically released: the MR-Linac (1.5 T) [88], MRI-Linac (1.0 T) [89] and Linac-MR (0.5 T) [90]. The MR-IGRT systems with higher field strengths should enable more flexible

imaging. However, when a magnetic field is present, the secondary electrons that are created during irradiation will have different paths due to the Lorentz forces than when no magnetic field is present. This effect is magnetic field strength dependent and therefore is stronger for the MR-Linac (1.5 T) than the Viewray (0.35 T) [91].

MR-IGRT allows patient alignment based on MRI images, in which soft tissue is clearly visible. Furthermore, MR-IGRT systems can obtain fast 2D images for 2D tracking/gating of the tumor during irradiation [92]. Additionally, MR-IGRT systems can acquire 4DMRI before treatment [77, 82, 85]. When continuously interleaving the acquisition of two perpendicular oriented 2D images continuously during treatment, the 4DMRI can be used as a model to enable 3D tumor tracking [93].

With the introduction of MR-IGRT, the added value of intratumoral fiducial markers will become debatable. MRI images allow for soft tissue matches, making markers potentially redundant. However, we believe that if markers are visible on MRI, they could still serve for several purposes in such systems. During radiotherapy treatment, it is unlikely that daily (typically 30 days) intravenous CE MRI will be obtained due to patient discomfort (due to intravenous infusion) and associated costs. Some tumors may be poorly visible on other sequences. Therefore, soft tissue matches may be based on several landmarks close to the tumor. It was shown that landmarks such as biliary stents or bony anatomy do not necessary correlate to tumor position due to deformations between both acquisitions [38, 41]. Therefore, intratumoral markers may help to align patients or to verify soft tissue registrations.

Then again, one might argue that with markers visible on MRI, CT and CBCT, there is limited added value in introducing an expensive MR-IGRT system compared to a conventional linac. To align patients, intratumoral markers can be used for conventional linacs. Furthermore, the markers can be utilized for 2D tracking of the tumor [94]. Moreover, 3D tracking of fiducial markers can be achieved by linacs with two integrated X-ray tubes at an angle to each other, such as in the Vero [95, 96]. The 3D position of the marker is then found by triangulation. For conventional linacs 3D tracking can be achieved using a combination of triangulation and/or digital tomosynthesis. We already illustrated the feasibility of tracking bony anatomy in 3D during treatment for this approach [97, 98].

Despite the possibilities of conventional linacs and markers, there are some clear advantages of MR-IGRT systems over conventional linacs. MR-IGRT may not require implanting of fiducial markers in the future. This would reduce the related work load and increase patient comfort. Also, contrary to the MR-IGRT, the conventional linac will require several CBCTs during treatment. These CBCTs cause additional radiation exposure for a large region of the patient. Finally, one of the major advantages of MR-IGRT, compared to a conventional IGRT, is that for several tumors, MR-IGRT provides

sufficient soft tissue contrast and tumor visibility to see organs at risk (OARs) and tumor prior to and during treatment. This enables daily replanning [99-101] or real-time replanning during irradiation of such tumors, which can increase the irradiation dose for the tumor and decrease dose for OARs and other healthy tissue. This is particularly interesting for tumors in a deformable anatomy and for tumors that change shape during treatment.

References

1. van Herk M. Errors and margins in radiotherapy. *Semin. Radiat. Oncol.* 2004; 14(1):52–64.
2. van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2000; 47(4):1121–1135.
3. Nijkamp J, de Haas-Kock DFM, Beukema JC et al. Target volume delineation variation in radiotherapy for early stage rectal cancer in the Netherlands. *Radiother. Oncol.* 2012; 102(1):14–21.
4. Nielsen MH, Berg M, Pedersen AN et al. Delineation of target volumes and organs at risk in adjuvant radiotherapy of early breast cancer: National guidelines and contouring atlas by the Danish breast cancer cooperative group. *Acta Oncol. (Madr).* 2013; 52(1):703–710.
5. Rasch CRN, Steenbakkers RJHM, Fitton I et al. Decreased 3D observer variation with matched CT-MRI, for target delineation in nasopharynx cancer. *Radiat. Oncol.* 2010; 5:21.
6. Rasch C, Keus R, Pameijer FA et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 1997; 39(4):841–848.
7. Villeirs GM, van Vaerenbergh K, Vakaet L et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther. Onkol.* 2005; 181(7):424–430.
8. Rasch C, Barillot I, Remeijer P et al. Definition of the prostate in CT and MRI: A multi-observer study. *Int. J. Radiat. Oncol. Biol. Phys.* 1999; 43(1):57–66.
9. den Hartogh MD, Philippons MEP, van Dam IE et al. MRI and CT imaging for preoperative target volume delineation in breast-conserving therapy. *Radiat. Oncol.* 2014; 9(1):63.
10. Dalah E, Moraru I, Paulson E et al. Variability of target and normal structure delineation using multimodality imaging for radiation therapy of pancreatic cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):633–640.
11. Li J, Denniston KA, Hussain SM et al. Comparison of CT and MRI-based gross tumor volume and organ at risk delineation for pancreatic cancer patients undergoing neoadjuvant stereotactic body radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 90(1):S356.
12. Hall WA, Mikell JL, Mittal P et al. Tumor size on abdominal MRI versus pathologic specimen in resected pancreatic adenocarcinoma: Implications for radiation treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 86(1):102–107.
13. Stroom JC, Heijmen BJM. Geometrical uncertainties, radiotherapy planning margins, and the ICRU-62 report. *Radiother. Oncol.* 2002; 64(1):75–83.
14. Van de Moortele P, Pfeuffer J, Glover GH et al. Respiration-induced B0 fluctuations and their spatial distribution in the human brain at 7 Tesla. *Magn. Reson. Med.* 2002; 47(5):888–895.
15. Moerland MA, Beersma R, Bhagwandien R et al. Analysis and correction of geometric distortions in 1.5 T magnetic resonance images for use in radiotherapy treatment planning. *Phys. Med. Biol.* 1995; 40(10):1651–1664.
16. Rylander S, Thörnqvist S, Haack S et al. Intensity profile based measurement of prostate gold markers influence on 1.5 and 3T diffusion-weighted MR images. *Acta Oncol.* 2011; 50(6):866–872.
17. van der Laan MJ, Bartels LW, Bakker CJG et al. Suitability of 7 aortic stent-graft models for MRI-based surveillance. *J. Endovasc. Ther.* 2004; 11(4):366–371.
18. Adams GJ, Baltazar U, Karmonik C et al. Comparison of 15 different stents in superficial femoral arteries by high resolution MRI ex vivo and in vivo. *J. Magn. Reson. Imaging* 2005; 22(1):125–135.
19. Klemm T, Duda S, Machann J et al. MR imaging in the presence of vascular stents: A systematic assessment of artifacts for various stent orientations, sequence types, and field strengths. *J. Magn. Reson. Imaging* 2000; 12(4):606–615.
20. Kim Y, Muruganandham M, Modrick JM, Bayouth JE. Evaluation of artifacts and distortions of titanium applicators on 3.0-tesla MRI: Feasibility of titanium applicators in MRI-guided brachytherapy for gynecological cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2011; 80(3):947–955.

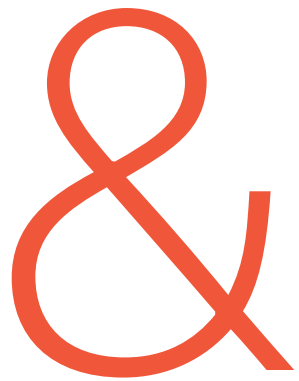
21. Hargreaves BA, Worters PW, Pauly KB et al. Metal-induced artifacts in MRI. *Am. J. Roentgenol.* 2011; 197(3):547–555.
22. Haack S, Nielsen SK, Lindegaard JC et al. Applicator reconstruction in MRI 3D image-based dose planning of brachytherapy for cervical cancer. *Radiother. Oncol.* 2009; 91(2):187–193.
23. van Heerden LE, Gurney-Champion OJ, van Kesteren Z et al. Quantification of image distortions on the Utrecht interstitial CT/MR brachytherapy applicator at 3T MRI. *Brachytherapy* 2015; 15(1):118–126.
24. Guo Y, Jiang X. Simulations of the stent artifacts in magnetic resonance imaging. *IEEE Trans. Magn.* 2012; 48(2):659–662.
25. Jonsson JH, Garpebring A, Karlsson MG, Nyholm T. Internal fiducial markers and susceptibility effects in MRI - Simulation and measurement of spatial accuracy. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 82(5):1612–1618.
26. Benoit-Cattin H, Collewet G, Belaroussi B et al. The SIMRI project: a versatile and interactive MRI simulator. *J. Magn. Reson.* 2005; 173(1):97–115.
27. Lagerburg V, Moerland MA, Seppenwoolde JH et al. Simulation of the artefact of an iodine seed placed at the needle tip in MRI-guided prostate brachytherapy. *Phys. Med. Biol.* 2008; 53(5):N59–N67.
28. Yoder DA, Zhao Y, Paschal CB, Fitzpatrick JM. MRI simulator with object-specific field map calculations. *Magn. Reson. Imaging* 2004; 22(3):315–328.
29. Cusack R, Brett M, Osswald K. An evaluation of the use of magnetic field maps to undistort echo-planar images. *Neuroimage* 2003; 18(1):127–142.
30. Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from B0 field variations. *Magn. Reson. Med.* 1995; 34(1):65–73.
31. Boegle R, MacLaren J, Zaitsev M. Combining prospective motion correction and distortion correction for EPI: Towards a comprehensive correction of motion and susceptibility-induced artifacts. *Magn. Reson. Mater. Physics, Biol. Med.* 2010; 23(4):263–273.
32. Li Y, Xu N, Fitzpatrick JM, Dawant BM. Geometric distortion correction for echo planar images using nonrigid registration with spatially varying scale. *Magn. Reson. Imaging* 2008; 26(10):1388–1397.
33. In MH, Posnansky O, Beall EB et al. Distortion correction in EPI using an extended PSF method with a reversed phase gradient approach. *PLoS One* 2015; 10(2):e0116320.
34. van Buuren LD, Polders D, Milder MT et al. Correcting diffusion weighted MR images for signal pile-up and distortions near gas pockets. *Proc. Intl. Soc. Mag. Reson. Med.* 24, 2016:2077.
35. Crijns SPM, Bakker CJG, Seevinck PR et al. Towards inherently distortion-free MR images for image-guided radiotherapy on an MRI accelerator. *Phys. Med. Biol.* 2012; 57(5):1349–1358.
36. Guckenberger M, Baier K, Guenther I et al. Reliability of the bony anatomy in image-guided stereotactic radiotherapy of brain metastases. *Int. J. Radiat. Oncol. Biol. Phys.* 2007; 69(1):294–301.
37. Porter BC, Rubens DJ, Strang JG et al. Three-dimensional registration and fusion of ultrasound and MRI using major vessels as fiducial markers. *IEEE Trans. Med. Imaging* 2001; 20(4):354–359.
38. van der Horst A, Wognum S, Dávila Fajardo R et al. Interfractional position variation of pancreatic tumors quantified using intratumoral fiducial markers and daily cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 87(1):202–208.
39. Nederveen AJ, Dehnad H, van der Heide UA et al. Comparison of megavoltage position verification for prostate irradiation based on bony anatomy and implanted fiducials. *Radiother. Oncol.* 2003; 68(1):81–88.
40. Nederveen AJ, van der Heide UA, Dehnad H et al. Measurements and clinical consequences of prostate motion during a radiotherapy fraction. *Int. J. Radiat. Oncol. Biol. Phys.* 2002; 53(1):206–214.
41. van der Horst A, Lens E, Wognum S et al. Limited role for biliary stent as surrogate fiducial marker in pancreatic cancer: Stent and intratumoral fiducials compared. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 89(3):641–648.

42. Huisman HJ, Fütterer JJ, van Lin ENJT et al. Prostate cancer: Precision of integrating functional MR imaging with radiation therapy treatment by using fiducial gold markers. *Radiology* 2005; 236(1):311–317.
43. Kapanen M, Collan J, Beule A et al. Commissioning of MRI-only based treatment planning procedure for external beam radiotherapy of prostate. *Magn. Reson. Med.* 2013; 70(1):127–135.
44. Nichol AM, Brock KK, Lockwood GA et al. A magnetic resonance imaging study of prostate deformation relative to implanted gold fiducial markers. *Int. J. Radiat. Oncol. Biol. Phys.* 2007; 67(1):48–56.
45. Chan MF, Cohen G ad N, Deasy JO. Qualitative evaluation of fiducial markers for radiotherapy imaging. *Technol. Cancer Res. Treat.* 2014; 14(3):1–7.
46. de Souza Lawrence L, Ford E, Gilbert C et al. Novel applications of an injectable radiopaque hydrogel tissue marker for management of thoracic malignancies. *Chest* 2013; 143(6):1635–1641.
47. Bair RJ, Bair E, Viswanathan AN. Radiopaque polymer hydrogel used as a fiducial marker in gynecologic brachytherapy. *Brachytherapy* 2014; 13(6):S73.
48. Terry JL, Milovancev M, Nemanic S. In vitro evaluation of a novel fiducial marker for computed tomography and magnetic resonance imaging of soft tissues in small animals. *Am. J. Vet. Res.* 2014; 75(11):974–981.
49. Dávila Fajardo R, Lekkerkerker SJ, van der Horst A et al. EUS-guided fiducial markers placement with a 22-gauge needle for image-guided radiation therapy in pancreatic cancer. *Gastrointest. Endosc.* 2014; 79(5):851–855.
50. Gaeta M, Benedetto C, Minutoli F et al. Use of diffusion-weighted, intravoxel incoherent motion, and dynamic contrast-enhanced MR imaging in the assessment of response to radiotherapy of lytic bone metastases from breast cancer. *Acad. Radiol.* 2014; 21(10):1286–1293.
51. Koh DM. Science to practice: Can intravoxel incoherent motion diffusion-weighted MR imaging be used to assess tumor response to antivascular drugs? *Radiology* 2014; 272(2):307–308.
52. van der Heide UA, Houweling AC, Groenendaal G et al. Functional MRI for radiotherapy dose painting. *Magn. Reson. Imaging* 2012; 30(9):1216–1223.
53. Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol.* 2005; 6(2):112–117.
54. Thorwarth D, Eschmann SM, Paulsen F, Alber M. Hypoxia dose painting by numbers: A planning study. *Int. J. Radiat. Oncol. Biol. Phys.* 2007; 68(1):291–300.
55. Barral M, Taouli B, Guio B et al. Diffusion-weighted MR imaging of the pancreas: Current status and recommendations. *Radiology* 2015; 274(1):45–63.
56. Nguyen C, Fan Z, Sharif B et al. In vivo three-dimensional high resolution cardiac diffusion-weighted MRI: A motion compensated diffusion-prepared balanced steady-state free precession approach. *Magn. Reson. Med.* 2014; 72(5):1257–1267.
57. Xie Y, Yu W, Fan Z et al. High resolution 3D diffusion MRI of carotid plaque on 3T. *J. Cardiovasc. Magn. Reson.* 2014; 16(Suppl 1):O28.
58. Sinha U, Sinha S. High speed diffusion imaging in the presence of eddy currents. *J. Magn. Reson. Imaging* 1996; 6(4):657–666.
59. Zhang Q, Coolen BF, Strijkers GJ et al. High-resolution distortion-free diffusion imaging of the prostate using stimulated echo based turbo spin echo (DPsti-TSE) sequence. *Proc. Intl. Soc. Mag. Reson. Med.* 24, Singapore: 2016:577.
60. McNab JA, Miller KL. Steady-state diffusion-weighted imaging: theory, acquisition and analysis. *NMR Biomed.* 2010; 23(7):781–793.
61. Bieri O, Ganter C, Scheffler K. Quantitative in vivo diffusion imaging of cartilage using double echo steady-state free precession. *Magn. Reson. Med.* 2012; 68(3):720–729.
62. Lee Y, Lee SS, Kim N et al. Intravoxel incoherent motion diffusion-weighted MR imaging of the liver: effect of triggering methods on regional variability and measurement repeatability of quantitative parameters. *Radiology* 2015; 196(6):405–415.
63. Taouli B, Koh DM. Diffusion-weighted MR imaging of the liver. *Radiology* 2010; 254(1):47–66.

64. Taouli B, Thakur RK, Mannelli L et al. Renal lesions: Characterization with diffusion-weighted imaging versus contrast-enhanced MR imaging. *Radiology* 2009; 251(2):398–407.
65. Koh DM, Collins DJ, Orton MR. Intravoxel incoherent motion in body diffusion-weighted MRI: Reality and challenges. *Am. J. Roentgenol.* 2011; 196(6):1351–1361.
66. Le Bihan D, Breton E, Lallemand D et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 1988; 168(2):497–505.
67. Heusch P, Wittsack HJ, Heusner T et al. Correlation of biexponential diffusion parameters with arterial spin-labeling perfusion MRI: Results in transplanted kidneys. *Invest. Radiol.* 2013; 48(3):140–144.
68. Lemke A, Laun FB, Simon D et al. An in vivo verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. *Magn. Reson. Med.* 2010; 64(6):1580–1585.
69. Bennett KM, Schmainda KM, Bennett RT et al. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* 2003; 50(4):727–734.
70. van der Bel R, Gurney-Champion OJ, Froeling M et al. SP224TRI–Exponential approach for intravoxel incoherent motion analysis of multi b-value diffusion weighted MRI data follows GFR changes in healthy humans. *Nephrol. Dial. Transplant.* 2016; 31(suppl 1):i161.
71. Gurney-Champion OJ, van der Bel R, Froeling M et al. Diffusion parameters derived from multi b-value DWI-data as surrogate marker for kidney function. *Proc. Intl. Soc. Mag. Reson. Med.* 24, Singapore: 2016:3439.
72. Winfield JM, DeSouza NM, Priest AN et al. Modelling DW-MRI data from primary and metastatic ovarian tumours. *Eur. Radiol.* 2015; 25(7):2033–2040.
73. Jerome NP, Miyazaki K, Collins DJ et al. Repeatability of derived parameters from histograms following non-Gaussian diffusion modelling of diffusion-weighted imaging in a paediatric oncological cohort. *Eur. Radiol.* 2016. doi:10.1007/s00330-016-4318-2.
74. Orton MR, Messiou C, Collins D et al. Diffusion-weighted MR imaging of metastatic abdominal and pelvic tumours is sensitive to early changes induced by a VEGF inhibitor using alternative diffusion attenuation models. *Eur. Radiol.* 2016; 26(5):1412–1419.
75. Winfield JM, Orton MR, Collins DJ et al. Separation of type and grade in cervical tumours using non-mono-exponential models of diffusion-weighted MRI. *Eur. Radiol.* 2016. doi:10.1007/s00330-016-4417-0.
76. Wetscherek A, Stieltjes B, Laun FB. Flow-compensated intravoxel incoherent motion diffusion imaging. *Magn. Reson. Med.* 2015; 74(2):410–419.
77. van Kesteren Z, Tekelenburg D, Gurney-Champion OJ et al. Reconstruction and validation of T2-weighted 4D magnetic resonance imaging for radiotherapy treatment planning. *Proc. Intl. Soc. Mag. Reson. Med.* 24, Singapore: 2016:1118.
78. Du D, Caruthers SD, Glide-Hurst C et al. High-quality T2-weighted 4-dimensional magnetic resonance imaging for radiation therapy applications. *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 92(2):430–437.
79. Hu Y, Caruthers SD, Low DA et al. Respiratory Amplitude Guided 4-Dimensional Magnetic Resonance Imaging. *Int. J. Radiat. Oncol. Biol. Phys.* 2013; 86(1):198–204.
80. Winkelmann S, Schaeffter T, Koehler T et al. An optimal radial profile order based on the Golden Ratio for time-resolved MRI. *IEEE Trans. Med. Imaging* 2007; 26(1):68–76.
81. Wundrak S, Paul J, Ulrici J et al. A small surrogate for the golden angle in time-resolved radial MRI based on generalized fibonacci sequences. *IEEE Trans. Med. Imaging* 2015; 34(6):1262–1269.
82. Stemkens B, Tijssen RHN, De Senneville BD et al. Optimizing 4-dimensional magnetic resonance imaging data sampling for respiratory motion analysis of pancreatic tumors. *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 91(3):571–578.
83. Feng L, Grimm R, Block KT et al. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magn. Reson. Med.* 2014; 72(3):707–717.

84. Rank CM, Heußer T, Buzan MTA et al. 4D respiratory motion-compensated image reconstruction of free-breathing radial MR data with very high undersampling. *Magn. Reson. Med.* 2016. doi:10.1002/mrm.26206.
85. Schoormans J, Gurney-Champion O, Klaassen R et al. 4D radial fat-suppressed alternating-TR bSSFP MRI with compressed sensing reconstruction for abdominal imaging during free breathing. *Proc. Intl. Soc. Mag. Reson. Med.* 24, Singapore: 2016:1097.
86. Rank CM, Sauppe S, Heußer T et al. Five-dimensional respiratory and cardiac motion compensation based on strongly undersampled MR data. *Proc. Intl. Soc. Mag. Reson. Med.* 24, 2016:786.
87. Dempsey JF, Benoit D, Fitzsimmons JR et al. A device for realtime 3D image-guided IMRT. *Int. J. Radiat. Oncol. Biol. Phys.* 2005; 63(2):S202.
88. Raaymakers BW, Lagendijk JJW, Overweg J et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: Proof of concept. *Phys. Med. Biol.* 2009; 54(12):N229–N237.
89. Constantin DE, Holloway L, Keall PJ, Fahrig R. A novel electron gun for inline MRI-linac configurations. *Med. Phys.* 2014; 41(2):22301.
90. Fallone BG, Murray B, Rathee S et al. First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system. *Med. Phys.* 2009; 36(6):2084–2088.
91. Raaijmakers AJE, Raaymakers BW, Lagendijk JJW. Magnetic-field-induced dose effects in MR-guided radiotherapy systems: dependence on the magnetic field strength. *Phys. Med. Biol.* 2008; 53(4):909–923.
92. Kashani R, Tanderup K, Victoria JR et al. Feasibility of gating using a magnetic-resonance image guided radiation therapy (MR-IGRT) system. *Int. J. Radiat. Oncol. Biol. Phys.* 2014; 90(1):S843–S844.
93. Stermkens B, Tijssen RHN, de Senneville BD et al. Image-driven, model-based 3D abdominal motion estimation for MR-guided radiotherapy. *Phys. Med. Biol.* 2016; 61(14):5335–5355.
94. Lens E, van der Horst A, Kroon PS et al. Differences in respiratory-induced pancreatic tumor motion between 4D treatment planning CT and daily cone beam CT, measured using intratumoral fiducials. *Acta Oncol.* 2014; 53(9):1257–1264.
95. Depuydt T, Poels K, Verellen D et al. Treating patients with real-time tumor tracking using the Vero gimbaled linac system: Implementation and first review. *Radiother. Oncol.* 2014; 112(3):343–351.
96. Solberg TD, Medin PM, Ramirez E et al. Commissioning and initial stereotactic ablative radiotherapy experience with Vero. *J. Appl. Clin. Med. Phys.* 2014; 15(2):4685.
97. Verbakel W, Gurney-Champion OJ, Cuijpers JP et al. Digital tomosynthesis performance for spine tracking. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 84(3):S203–S204.
98. Gurney-Champion OJ, Dahele M, Mostafavi H et al. Digital tomosynthesis for verifying spine position during radiotherapy: a phantom study. *Phys. Med. Biol.* 2013; 58(16):5717–5733.
99. Vestergaard A, Hafeez S, Muren LP et al. The potential of MRI-guided online adaptive re-optimisation in radiotherapy of urinary bladder cancer. *Radiother. Oncol.* 2016; 118(1):154–159.
100. Kontaxis C, Bol GH, Lagendijk JJW, Raaymakers BW. Towards adaptive IMRT sequencing for the MR-linac. *Phys. Med. Biol.* 2015; 60(6):2493–2509.
101. Kerkhof EM, Raaymakers BW, van der Heide UA et al. Online MRI guidance for healthy tissue sparing in patients with cervical cancer: An IMRT planning study. *Radiother. Oncol.* 2008; 88(2):241–249.





Summary

Pancreatic cancer is one of the deadliest of all major cancers with a median survival of 4.4 months after diagnosis. The only known potentially curative treatment for pancreatic cancer is surgical resection. However, even after successful resection, patients still have a poor median survival of 23 months. Therefore, surgery may be combined with (neo)adjuvant chemotherapy and/or radiotherapy. Still, there are many challenges in radiotherapy that should be overcome to improve the treatment outcome further.

During radiotherapy, tumors are irradiated with ionizing radiation. This ionizing radiation can kill both tumor cells as well as healthy cells. Consequently, the goal of radiotherapy is to deliver the highest possible dose to the tumor, while limiting the dose to healthy tissue. To achieve this, a CT scan is obtained from the patient, on which the radiation oncologist delineates the target volume. The treatment is based on this delineation. However, for many types of tumors including pancreatic cancer, the tumor is poorly depicted on CT images. As a result, mistakes may occur during delineation. Such mistakes directly influence the treatment plan and can cause under dosage of the tumor or unnecessary irradiation to healthy tissue. In addition, some pancreatic cancers are less sensitive to irradiation. For patients with such tumors, radiotherapy will be less effective in the tumor, but will still cause damage to healthy tissue.

In this thesis, we investigated how MRI can be used to improve radiotherapy for pancreatic cancer patients. In the first part, the use of MRI to improve tumor delineation during treatment planning is discussed. In the second part, diffusion-weighted imaging (DWI) is optimized. This technique allows studying certain tissue properties as diffusivity and perfusion. Studying these properties potentially allows monitoring treatment response. Ultimately, the tissue properties measured before treatment can be used to predict treatment outcome. In both cases, unnecessary irradiation of patients with non-responding tumors can be prevented.

In **chapter 2** we assess the current state of delineating pancreatic tumors on CT by radiation oncologists (observers). Such delineations are used as a basis for all radiotherapy sessions and should be as accurate as possible. In this study, eight radiation oncologists delineated the gross tumor volume (GTV) on CT for four pancreatic cancer patients. We found that there was a large interobserver variability. There was poor overlap between the delineated volumes of the different observers. The local observer variation was largest close to pathological lymph nodes and biliary stents. These large variations were due to ambiguous instructions and poor protocol compliance. These results suggest that the delineated volumes on CT are

inaccurate and that the treatment may depend on the radiation oncologist involved in the treatment.

In **chapter 3** we show that by offering unmatched MRI images alongside the CT images, the observer variability decreases. The same eight radiation oncologists again delineated the GTV of the same four pancreatic cancer patients, now with MRI images available. The delineated volumes were significantly smaller (32% on average) when MRI was made available. Furthermore, the overall observer variation decreased by 35%. The decrease in interobserver variation suggests that more accurate delineations can be achieved by adding MRI.

In our study, CT and MRI images were not matched. In the future, matching MRI images to CT could help further decrease interobserver variability. To match MRI to CT, visible intratumoral fiducial markers are desirable. It is known that such markers are visible on CT and cone beam CT (CBCT). However, little is known about their visibility on MRI images. Therefore, in **chapter 4** the visibility and signal shift artifacts (susceptibility artifacts) caused by fiducial markers on MRI images is quantified using a sequence-independent approach. The signal shift artifact is an unwanted deformation in the MRI images due to inhomogeneities in the main magnetic field. In this sequence independent approach, visibility was assessed by measuring $T2^*$ relaxation times whereas signal shift artifacts were evaluated by measuring inhomogeneities in the main magnetic field (ΔB_0). Using phantom measurements, we showed that for the ten radiopaque gold fiducial markers measured, there is a strong correlation between $T2^*$ and $|\Delta B_0|$, and hence, more visible markers also show larger artifacts. We also illustrated that the marker that was best visible in the phantom, was also visible *in vivo*, in a pancreatic cancer patient. However, this marker also caused artifacts on DWI images. Furthermore, we showed that the marker with the smallest variation in $|\Delta B_0|$ showed no artifacts in DWI images, but also was not visible *in vivo*. As a result, it is now possible to select markers that best fit the application. One can select markers that are visible on MRI, allowing for MRI-CT or MRI-CBCT matching. One can also select a marker that allow for artifact-free MRI, which is relevant for DWI images.

Seventy percent of pancreatic cancer patients that receive radiotherapy have a biliary stent which prevents the tumor from blocking the biliary duct. Therefore, in **chapter 5** the artifacts of eight commonly used biliary stents are quantified using a similar sequence-independent approach. The eight stents, made of nitinol, platinum-cored nitinol, stainless steel or polyethylene, were placed in a phantom, and MRI images were acquired. It was shown that the susceptibility artifacts ($T2^*$ and ΔB_0) dominate over the gradient-induced artifacts and radiofrequency (RF) induced artifacts. The susceptibility artifacts grow with field strength and were larger at 3 T than at 1.5 T. The susceptibility artifacts vary as a function of the angle between the

stent and the main magnetic field. They were largest when the stent was oriented perpendicular to the main magnetic field and were smallest at parallel orientation. The stent's material also influences the artifact size: the largest susceptibility artifacts were found in the stainless steel stent, whereas the smallest were found in a polyethylene stent. Finally, it was illustrated that the effect of the stent on $T2^*$ and $\Delta B0$ in a phantom were similar to the effect *in vivo*.

In **chapter 6** we optimized the settings of the alternating repetition time balanced steady state free precession (ATR-SSFP) sequence for abdominal imaging in a single breath-hold. At these optimal settings, ATR-SSFP has a $T2$ -like contrast. This contrast can be useful for distinguishing between tissue types. Two radiologists scored ATR-SSFP images as well as images obtained by the more easily implemented conventional fat-saturated balanced steady state free precession (FS-bSSFP) of fifteen pancreatic cancer patients. From the radiologists' scores, it was shown that at optimal settings, the ATR-SSFP was preferred. With its $T2$ -like contrast, the high isotropic resolution and its inherent fat saturation, ATR-SSFP imaging is a strong alternative to conventional $T2$ -weighted imaging, especially for radiotherapy.

The intravoxel incoherent motion (IVIM) model for DWI is a promising model to monitor and predict treatment response. For this purpose, measurements are required to be accurate and precise. However, the acquisition and post-processing of the DWI data have not been optimized for pancreatic imaging. Therefore, in **Chapter 7** we determined the combination of b -values and signal averages for DWI that rendered the minimum acquisition time necessary for accurate and precise IVIM modeling. We assessed the repeatability (measure for precision) and systematic errors (measure for accuracy) of IVIM model fit parameters in sixteen volunteers as a function of the b -values and averages used in the IVIM model fit. We showed that acquisition could be done in 5:15 minutes in the pancreas and in 2:15 minutes for the liver.

In **Chapter 8** the performance of six IVIM model fit algorithms and seven DWI fit models was assessed in DWI data from fourteen pancreatic cancer patients. The Levenberg-Marquardt least squares fit with fixed pseudo-diffusion coefficient, and the Bayesian fit scored best out of the IVIM model fit algorithms. Although, the IVIM model scored best overall, other DWI models may be preferred for specific purposes. For example, for depicting pancreatic tumors, the perfusion fraction f_2 from the triexponential model is most promising. For treatment monitoring, the stretched exponent model is preferred due to its high precision in both fit parameters.

In **chapter 9**, the work is summarized and contextualized to a larger perspective. In addition, two techniques are discussed: 4DMRI and MRI-guided radiotherapy. 4DMRI is a set of MRI images showing the patient's anatomy throughout a typical respiratory cycle. MRI-guided radiotherapy has recently gained much momentum

with the development of radiotherapy treatment systems with integrated MRI. Such systems allow monitoring both tumor as well as organs at risk during radiotherapy treatment.

In conclusion, MRI should become part of the standard clinical routine in radiotherapy of pancreatic cancer patients. We showed the improvement of tumor delineations with the addition of MRI images. We can use fiducial markers that are visible on MRI, enabling marker based image registration. However, one should be aware of potential artifacts from visible markers. Furthermore, the potential artifacts caused by biliary stents are well understood and can be minimized by choosing appropriate alternative stents. The ATR-SSFP sequence can be used to obtain high-resolution images with T_2 -like contrast and fat saturation. Also, we optimized DWI acquisitions and post-processing in pancreatic cancer patients.

Samenvatting

Met een mediane overleving van 4,4 maanden behoort alvleesklierkanker tot een van de meest dodelijke kankersoorten. De enige potentieel genezende behandeling voor patiënten met alvleesklierkanker is een resectie. Maar zelfs na een succesvolle resectie hebben deze patiënten nog steeds een vrij korte mediane overleving van 23 maanden. Daarom wordt een resectie vaak gecombineerd met (neo)adjuvante chemotherapie en/of radiotherapie. Er bestaan echter nog veel uitdagingen voor het toepassen van radiotherapie in patiënten met alvleesklierkanker.

Tijdens radiotherapie worden tumoren bestraald met ioniserende straling. Deze straling is schadelijk voor tumorcellen, maar ook voor gezonde cellen. Daarom is het doel van radiotherapie om een zo hoog mogelijke stralingsdosis te geven aan de tumor, met een beperkte stralingsdosis in gezond weefsel. Hiervoor wordt voorafgaande aan de behandeling vaak een computertomografie (CT) scan gemaakt van de patiënt. Op deze CT-scan definieert de radiotherapeut het te bestralen doelgebied (intekening). Deze intekening wordt gebruikt om het bestralingsplan te genereren. Echter, voor veel tumorsoorten, waaronder alvleeskliertumoren, is de tumor slecht zichtbaar op CT-beelden. Daardoor kunnen er vergissingen optreden tijdens de intekening. Zulke vergissingen hebben gevolgen voor het bestralingsplan. Delen van de tumor kunnen daardoor een lagere dosis krijgen dan wenselijk is en gezond weefsel kan onnodig worden bestraald. Daarbij komt dat niet alle alvleeskliertumoren even gevoelig zijn voor ioniserende straling waardoor radiotherapie niet voor alle patiënten effectief is.

In dit proefschrift is onderzocht hoe MRI kan worden ingezet om radiotherapie voor patiënten met alvleesklierkanker te bevorderen. In het eerste deel wordt besproken hoe MRI kan helpen tijdens het definiëren van de te bestralen doelgebieden. In het tweede deel wordt diffusie gewogen MRI (DWI) geoptimaliseerd. Deze beeldvormende techniek maakt het meten van bepaalde weefseleigenschappen mogelijk zoals de mate van diffusie en perfusie. Dit kan helpen om de reactie van de tumor op radiotherapie te bestuderen en daarmee de effectiviteit van radiotherapie te bepalen. Potentieel kan de effectiviteit van radiotherapie zelfs voor de behandeling al worden voorspeld aan de hand van de weefseleigenschappen. In beide gevallen kan onnodige bestraling worden voorkomen.

In **hoofdstuk 2** wordt het huidige niveau van pancreastumor intekeningen op CT bepaald. Zulke intekeningen zijn de basis voor veel radiotherapie behandelingen en het is dus wenselijk dat deze van hoog niveau zijn. Acht radiotherapeuten hebben voor deze studie het tumorvolume (gross tumor volume; GTV) ingetekend op een CT-scan van vier patiënten met alvleesklierkanker. We vonden een grote variatie van intekeningen (intekenvariatie) tussen de radiotherapeuten. Ook was er een beperkte overlap tussen de intekeningen van de verschillende radiotherapeuten. De grootste

lokale intekenvariaties waren dichtbij pathologische lymfeklieren en stents. Deze variaties zijn het gevolg van enerzijds onvolledige instructies en anderzijds het niet opvolgen van instructies door de deelnemende artsen. De studie suggereert dat de huidige tekeningen op CT inaccuraat zijn. De radiotherapiebehandeling kan dus sterk afhangen van de behandelende radiotherapeut.

In **hoofdstuk 3** laten we zien dat de intekenvariatie vermindert als er naast de CT-beelden ook MRI-beelden worden aangeboden aan de radiotherapeut. In deze studie hebben dezelfde acht radiotherapeuten het GTV nogmaals ingetekend op CT-beelden van dezelfde vier patiënten, ditmaal met de beschikking over MRI-beelden. De volumes van de ingetekende GTVs waren significant kleiner (gemiddeld 32%) in de studie waar MRI beschikbaar was dan in de studie waar MRI niet beschikbaar was. De totale intekenvariatie nam af met 35% toen MRI beschikbaar was. Dit suggereert dat de tekeningen accurater zijn als MRI-beelden beschikbaar zijn tijdens het intekenen.

In ons onderzoek waren de CT- en MRI-beelden niet gematcht. In het vervolg zou matchen van de MRI-beelden aan de CT-beelden kunnen worden toegepast om de tekeningen verder te verbeteren. Hiervoor is het wenselijk dat er markers in de tumor worden geplaatst die zichtbaar zijn op beide beeldvormende modaliteiten. Zulke markers kunnen dan als leidraad dienen tijdens het matchen. Het is aangetoond dat zulke markers goed zichtbaar zijn op CT en cone beam CT (CBCT), maar over de zichtbaarheid op MRI was weinig bekend. In **hoofdstuk 4** hebben we daarom een sequentie-onafhankelijke aanpak ontwikkeld om de zichtbaarheid van markers en de signaalverschuivingsartefacten die door markers worden veroorzaakt, te kwantificeren. Een signaalverschuivingsartefact is de ongewenste vervorming van het MRI-beeld als gevolg van inhomogeniteiten in het magnetische veld. In deze aanpak wordt de zichtbaarheid geëvalueerd aan de hand van metingen van $T2^*$ relaxatietijd terwijl de signaalverschuivingsartefacten worden geëvalueerd aan de hand van metingen van inhomogeniteiten in het magnetisch veld (ΔB_0). Voor dit onderzoek hebben we een fantoom ontwikkeld met tien markers die zichtbaar waren op CT. We laten zien dat er een sterk verband is tussen $T2^*$ en $|\Delta B_0|$. Dat betekent dat markers die beter zichtbaar zijn, ook grotere signaalverschuivingen creëren. We toonden aan dat de best zichtbare marker uit het fantoom ook zichtbaar was *in vivo*, in een patiënt met alvleesklierkanker. Deze marker veroorzaakt echter signaalverschuivingsartefacten op DWI-beelden. Verder toonden we aan dat de marker met de kleinste variaties in $|\Delta B_0|$ geen signaalverschuivingsartefacten creëert in DWI-beelden. Deze marker is echter niet zichtbaar *in vivo*. Als gevolg van dit onderzoek kan men nu een marker selecteren die het best past bij de toepassing van het MRI onderzoek. Men kan kiezen voor markers waarmee nauwkeurige MRI-CT en MRI-CBCT matches mogelijk zijn. Men

kan ook een marker selecteren die geen signaalverschuivingsartefacten veroorzaakt, wat relevant is voor DWI.

Zeventig procent van de alvleesklierkankerpatiënten die radiotherapie krijgen hebben een stent in de galwegen. Deze wordt geplaatst omdat de tumor de galwegen anders dichtduwt. Daarom zijn in **hoofdstuk 5** de artefacten van acht stents gekwantificeerd met een soortgelijke sequentie-onafhankelijke aanpak. Van een fantoom met de acht stents werden MRI-beelden gemaakt. Aan de hand van deze beelden toonden wij aan dat de susceptibiliteitsartefacten ($T2^*$ en $\Delta B0$) dominant zijn over de gradiënt geïnduceerde artefacten en de radiofrequentie (RF) geïnduceerde artefacten. De susceptibiliteitsartefacten variëren met magneetsterkte en waren groter in een 3 T MRI-scanner dan in een 1,5 T MRI-scanner. De susceptibiliteitsartefacten variëren met de hoek die de stent maakt met het magnetische veld. De artefacten waren het kleinst als de stent parallel aan het magnetische veld stond, terwijl de artefacten het grootst waren als de stent loodrecht op het magnetisch veld stond. Stentmateriaal heeft ook invloed op de grootte van het artefact, waar de stent van roestvrijstaal de grootste artefacten gaf terwijl de polyethyleen stent de kleinste artefacten gaf. Tot slot toonden wij aan dat het effect van de stent op $T2^*$ en $\Delta B0$ in een fantoom in overeenstemming was met het effect *in vivo*.

In **hoofdstuk 6** hebben we de instellingen van “alternating repetition time balanced steady state free precession” (ATR-SSFP) sequentie geoptimaliseerd voor beeldvorming van het abdomen tijdens het vasthouden van de adem. Met deze optimale instellingen heeft ATR-SSFP een $T2$ -achtig contrast. Dit contrast kan helpen bij het onderscheiden van verschillende typen weefsel. Twee radiologen beoordeelden zowel ATR-SSFP beelden als conventionele “fat-saturated steady state free precession” (FS-bSSFP) beelden van vijftien alvleesklierkankerpatiënten. Uit deze beoordelingen bleek dat de radiologen de ATR-SSFP beelden prefereerden boven de FS-bSSFP. De ATR-SSFP beelden zijn een aantrekkelijk alternatief voor conventionele $T2$ -gewogen beelden in radiotherapie, door de hoge resolutie, het $T2$ -achtige contrast en de vetsuppressie die ATR-SSFP biedt.

Het “intravoxel incoherent motion” (IVIM) model voor DWI is een veelbelovend model om response van tumoren op een behandeling te detecteren. Hiervoor moeten de metingen wel accuraat en precies zijn. De acquisitie en nabewerking van DWI-data voor het fitten van het IVIM-model is echter nog niet geoptimaliseerd voor alvleesklierkankerpatiënten. Daarom hebben we in **hoofdstuk 7** de ideale combinatie van diffusie wegingen (b-waarden) en het aantal acquisities dat nodig is voor IVIM-model fits bepaald voor de kortst mogelijke acquisitie die nog accurate en precieze parameters oplevert. Hiervoor hebben we voor zestien vrijwilligers de reproduceerbaarheid van de meting (maat voor precisie) en de systematische fout

(maat voor accuraatheid) van de IVIM-model fit-parameters bepaald als functie van het aantal b-waarden en hoeveelheid metingen per b-waarde die zijn meegenomen in de fits. We hebben aangetoond dat een acquisitie van minimaal 5:15 minuten nodig is om reproduceerbare IVIM-model fits te doen voor de alvleesklier en 2:15 minuten in de lever.

In **hoofdstuk 8** wordt in veertien patiënten met alvleesklierkanker de prestatie van zes IVIM-model fit algoritmes onderling vergeleken, zowel als zeven DWI fit modellen. De Levenberg-Marquardt kleinste-kwadratenmethode fit met vaste waarde voor de pseudo-diffusie coëfficiënt en de Bayesiaanse fit scoorde het hoogst van alle IVIM-model fit algoritmes. Het IVIM-model scoorde over het algemeen het hoogst, maar afhankelijk van het doel kunnen andere diffusie modellen (zoals een triexponentieel of stretched exponentieel model) de voorkeur hebben. Bijvoorbeeld, voor het afbeelden van alvleeskliertumoren is de perfusie fractie f_2 van het triexponentieel model het best. Voor het detecteren van een response van een tumor op een behandeling is het stretched exponent model het meest belovend met reproduceerbare parameters.

In **hoofdstuk 9** wordt al het werk samengevat en in een ruimere context geplaatst. In dit hoofdstuk worden ook 4DMRI en MRI-begeleide radiotherapie besproken. Een 4DMRI is een set van MRI-beelden die de anatomie van de patiënt weergeven gedurende de ademhalingscyclus. MRI-begeleide radiotherapie is in een versnelling gekomen met de ontwikkeling van radiotherapie bestralingsapparaten met geïntegreerde MRI-systemen. Zulke systemen stellen de gebruiker in staat om de beweging van de tumor, zowel als de gezonde organen, te volgen gedurende de behandeling.

Uit het werk in dit proefschrift concluderen wij dat MRI een deel uit zou moeten maken van de standaard klinische routine van patiënten met alvleesklierkanker. We hebben laten zien dat MRI bijdraagt aan betere tumor intekeningen. Verder kunnen we markers gebruiken om de MRI-beelden te matchen aan CT- en CBCT-beelden. In dat geval moet men zich wel bewust zijn van de artefacten die zulke markers creëren op DWI-beelden. De potentiële artefacten die stents veroorzaken op MRI kunnen worden geminimaliseerd door een geschikte stent te selecteren. De ATR-SSFP sequentie kan worden gebruikt om beelden te krijgen met hoge resolutie, T2-achtig contrast en vetsuppressie. Verder hebben we DWI-acquisities en nabewerking geoptimaliseerd voor patiënten met alvleesklierkanker.

PhD Portfolio

PhD Training

	year	ECTS
General courses		
Pubmed	2013	0.1
Basiscursus Regelgeving en Organisatie van Klinisch Onderzoek (BROK)	2013	0.9
Practical Biostatistics	2013	1.5
Project Management	2013	0.6
Oral Presentation (English)	2014	0.8
Scientific Writing for Publication (English)	2014	1.5
Specific courses		
In Vivo NMR (Wageningen)	2012	3.0
Pulse Programming, Philips (Utrecht)	2013	1.5
Bayesian Statistics in Mathematica (Amsterdam)	2013	0.4
Diffusion: What it means and how to measure it (ESMRMB; Split, Croatia)	2013	1.2
Oncology course OOA	2014	1.5
Seminars, workshops and master classes		
Weekly MRF meeting (MR physics science club, 6 oral presentations)	2012–2016	7
Some Radiotherapy Physics meetings (science club, 2 oral presentations)	2012–2016	3
OOA retreat, Renesse (poster)	2013	1.25
OOA retreat, Renesse (talk)	2014	1.25
(Inter)national conferences		
ISMRM meeting, Salt Lake City, USA	2013	1.75
ISMRM/ESMRMB joint meeting, Milan, Italy (poster)	2014	2.25
ISMRM meeting, Toronto, Canada (2 posters)	2015	2.75
ISMRM meeting, Singapore (2 e-posters)	2016	2.75
ISMRM Benelux chapter, Rotterdam	2012	0.25
ISMRM Benelux chapter, Maastricht	2013	0.25
ISMRM Benelux chapter, Ghent, Belgium (poster)	2014	0.25
ISMRM Benelux chapter, Eindhoven (poster)	2015	0.25
ESTRO meeting, Vienna, Austria (poster)	2014	1.25

Teaching**year ECTS****Lecturing**

Matlab course (assistant)	2014-2016	0.75
<i>In Vivo</i> NMR course (assistant during exercises)	2016	0.25

Supervising

Bachelor's student (2 months)	2014	0.75
Bachelor's student (2.5 months)	2015	0.75
Master's student (1 year)	2016	2.25

Other

Chair of the organizing committee ISMRM Benelux chapter	2014	6
Organizer of the MRF meeting (MR physics science club)	2013-2016	0.5

List of Publications

Gurney-Champion OJ, Dahele M, Mostafavi H, Slotman BJ, Verbakel W. Digital tomosynthesis for verifying spine position during radiotherapy: a phantom study. *Phys. Med. Biol.* 2013; 58(16):5717–5733.

Gurney-Champion OJ, Lens E, van der Horst A, Houweling AC, Klaassen R, van Hooft JE, Stoker J, van Tienhoven G, Nederveen AJ, Bel A. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med. Phys.* 2015; 42(5):2638–2647.

Verbakel WFAR, **Gurney-Champion OJ**, Slotman BJ, Dahele M. Sub-millimeter spine position monitoring for stereotactic body radiotherapy using offline digital tomosynthesis. *Radiother. Oncol.* 2015; 115(2):223–228.

van Heerden LE, **Gurney-Champion OJ**, van Kesteren Z, Houweling AC, Koedooder C, Rasch CRN, Pieters BR, Bel A. Quantification of image distortions on the Utrecht interstitial CT/MR brachytherapy applicator at 3T MRI. *Brachytherapy* 2015; 15(1):118–126.

Gurney-Champion OJ, Nederveen AJ, Klaassen R, Engelbrecht MR, Bel A, van Laarhoven HWM, Stoker J, Goncalves SI. Revisiting the potential of alternating repetition time balanced steady-state free precession imaging of the abdomen at 3 T. *Invest. Radiol.* 2016; 51(9):560–568.

Gurney-Champion OJ, Froeling M, Klaassen R, Runge JH, Bel A, van Laarhoven HWM, Stoker J, Nederveen AJ. Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas. *Invest. Radiol.* 2016; 51(4):211–220.

Gurney-Champion OJ, Bruins Slot T, Lens E, van der Horst A, Klaassen R, van Laarhoven HWM, van Tienhoven G, van Hooft JE, Nederveen AJ, Bel A. Quantitative assessment of biliary stent artifacts on MR images: Potential implications for target delineation in radiotherapy. *Med. Phys.* 2016; 43(10):5603–5615.

Lens E, **Gurney-Champion OJ**, Tekelenburg DR, van Kesteren Z, Parkes MJ, van Tienhoven G, Nederveen AJ, van der Horst A, Bel A. Abdominal organ motion during breath-holding from inhalation and exhalation: different lung volumes compared. *Radiother. Oncol.* 2016; In press

Gurney-Champion OJ, Klaassen R, Froeling M, Barbieri S, Stoker J, Engelbrecht MRW, Wilmink JW, Besselink MG, Bel A, van Laarhoven HWM, Nederveen AJ. How best to fit diffusion-weighted magnetic resonance imaging data of pancreatic cancer patients. Submitted

Gurney-Champion OJ, Versteijne E, van der Horst A, Lens E, Rütten H, Heerkens HD, Paardekooper GMRM, Berbee M, Rasch CRN, Stoker J, Engelbrecht MRW, van Herk M, Nederveen AJ, Klaassen R, van Laarhoven HWM, van Tienhoven G, Bel A. Addition of MRI improves interobserver variation in CT-based pancreatic tumor delineation. Submitted

Versteijne E*, **Gurney-Champion OJ***, van der Horst A, Lens E, Kolff MW, Buijsen J, Ebrahimi G, Neelis K, Rasch CRN, Stoker J, van Herk M, Bel A, van Tienhoven G. Considerable interobserver variation in delineation of pancreatic cancer on 3DCT and 4DCT: a multi-institutional study. Submitted

van der Bel R*, **Gurney-Champion OJ***, Froeling M, Stroes ESG, Nederveen AJ, Krediet CTP. A tri-exponential model for intravoxel incoherent motion analysis during pharmacological renal perfusion modulation in humans. Submitted

van der Bel R, Verbree J, **Gurney-Champion OJ**, van Osch MJP, Stroes ESG, Nederveen AJ, Krediet CTP. Sympathetic activation by Lower Body Negative Pressure does not affect kidney oxygenation in healthy men. Submitted

* both authors contributed equally to this work.

International Conference Proceedings

Verbakel W, **Gurney-Champion OJ**, Cuijpers JP, Slotman BJ, Senan S, Dahele M. Digital tomosynthesis performance for spine tracking. *Int. J. Radiat. Oncol. Biol. Phys.* 2012; 84:S203–S204.

Gurney-Champion OJ, Lens E, van der Horst A, Houweling AC, van Hooft JE, Stoker J, van Tienhoven G, Nederveen AJ, Bel A. Establishing adequately visible endoscopically implantable gold markers for CT, cone-beam CT and MRI. *Radiother. Oncol.* 2014; 111:S105–S106.

Gurney-Champion OJ, Froeling M, Klaassen R, Stoker J, van Tienhoven G, van Laarhoven HWM, Bel A, Nederveen AJ. Assessment of reproducibility of IVIM based perfusion fraction and diffusion coefficient in the pancreas. *Proc. Intl. Soc. Mag. Reson. Med.* 22. Milan; 2014:2243.

Klaassen R, **Gurney-Champion OJ**, ter Voert EGW, Heerschap A, Bijlsma MF, Besselink MGH, van Tienhoven G, Nio CY, Stoker J, Punt CJA, Wilmink JW, van Laarhoven HWM, Nederveen AJ. Motion correction of high temporal 3T dynamic contrast enhanced MRI of pancreatic cancer – Preliminary results. *Proc. Intl. Soc. Mag. Reson. Med.* 22.; 2014:3667.

van Heerden LE, **Gurney-Champion OJ**, Van Kesteren Z, Pieters BR, Bel A. OC-0093: Quantification of deformations on 3T MRI for the Utrecht Interstitial CT/MR brachytherapy applicator. *Radiother. Oncol.* 2015; 115:S47.

Gurney-Champion OJ, Froeling M, Klaassen R, van Laarhoven HWM, Stoker J, Bel A, Nederveen AJ. Intra-voxel incoherent motion modelling of diffusion weighted MRI data is feasible in 5 minutes scan time. *Proc. Intl. Soc. Mag. Reson. Med.* 23. Toronto; 2015:2868.

Gurney-Champion OJ, Klaassen R, Stoker J, Bel A, van Laarhoven HWM, Nederveen AJ, Goncalves, SI. Revisiting the potential of alternating repetition time balanced steady state free precession imaging in the abdomen at 3 T. *Proc. Intl. Soc. Mag. Reson. Med.* 23. Toronto; 2015:1523.

Klaassen R, Steins A, **Gurney-Champion OJ**, Bijlsma MF, Wijkstra H, van Tienhoven G, Besselink MGH, Wilmink JW, van de Vijver MJ, Stoker J, Nederveen AJ, van Laarhoven HWM. Correlating post-operative whole mount immunohistochemistry to functional MRI parameters in pancreatic cancer. *Proc. Intl. Soc. Mag. Reson. Med.* 23. Toronto; 2015:141.

Steins A, Klaassen R, **Gurney-Champion O**, Bijlsma M, Medema JP, Wijkstra H, van Tienhoven G, Busch O, Punt C, Besselink M, Wilmink H, van de Vijver M, Stoker J, Nederveen A, van Laarhoven H. Abstract 421: The role of the tumor microenvironment of pancreatic cancer to predict treatment outcome. *Cancer Res.* 2015; 75:421.

van Wieringen N, van Heerden L, **Gurney-Champion O**, van Kesteren Z, Houweling A, Pieters B, Bel A. SU-E-J-216: A sequence independent approach for quantification of MR image deformations from brachytherapy applicators. *Med. Phys.* 2015; 42:3315.

van der Bel R, **Gurney-Champion OJ**, Froeling M, Nederveen AJ, Krediet CTP. SP224TRI-Exponential approach for intravoxel incoherent motion analysis of multi b-value diffusion weighted MRI data follows GFR changes in healthy humans. *Nephrol. Dial. Transplant.* 2016; 31:i161.

Versteijne E, **Gurney-Champion O**, van der Horst A, Lens E, Kolff M, Heerkens H, Paardekooper G, Berbee M, Buijsen J, van de Putte P, Neelis K, van Herk M, Bel A, van Tienhoven G. PO-0710: Large interobserver variation of delineated target volumes of pancreatic cancer in the Netherlands. *Radiother. Oncol.* 2016; 119:S331-332.

Tekelenburg D, **Gurney-Champion O**, Lens E, van der Horst A, Nederveen AJ, Biegun AK, Bel A, van Kesteren Z. PO-0913: Clinically applicable T2-weighted 4D magnetic resonance imaging with good abdominal contrast. *Radiother. Oncol.* 2016; 119:S440-S441.

Gurney-Champion OJ, Klaassen R, Froeling M, Stoker J, Wilmink JW, Bel A, van Laarhoven HWM, Nederveen AJ. Comparison of six different diffusion weighted MRI models in pancreatic cancer patients. *Proc. Intl. Soc. Mag. Reson. Med.* 24. Singapore; 2016:3434.

Gurney-Champion OJ, van der Bel R, Froeling M, Krediet CTP, Nederveen AJ. Diffusion parameters derived from multi b-value DWI-data as surrogate marker for kidney function. Proc. Intl. Soc. Mag. Reson. Med. 24. Singapore; 2016:3439.

Schoormans J, **Gurney-Champion O**, Klaassen R, Runge JH, Gonçalves SI, Coolen BF, Motaal AG, van Laarhoven HWM, Stoker J, Nederveen AJ, Strijkers GJ. 4D radial fat-suppressed alternating-TR bSSFP MRI with compressed sensing reconstruction for abdominal Imaging during free breathing. Proc. Intl. Soc. Mag. Reson. Med. 24. Singapore; 2016:1097.

van Kesteren Z, Tekelenburg D, **Gurney-Champion O**, Nederveen A, Lens E, van der Horst A, Biegun A, Bel A. Reconstruction and validation of T2-weighted 4D magnetic resonance imaging for radiotherapy treatment planning. Proc. Intl. Soc. Mag. Reson. Med. 24. Singapore; 2016:1118.

Runge JH, Klaassen R, **Gurney-Champion OJ**, van Laarhoven HWM, Sinkus R, Nederveen AJ, Stoker J. A feasibility study to perform combined MR elastography, IVIM and DCE-MRI in pancreatic cancer patients. Proc. Intl. Soc. Mag. Reson. Med. 24. Singapore; 2016:1126.

Zhang Q, Coolen BF, Strijkers GJ, van Buuren L, van der Heide U, **Gurney-Champion OJ**, Gonçalves SI, Nederveen AJ. High-resolution distortion-free diffusion imaging of the prostate using stimulated echo based turbo spin echo (DPsti-TSE) sequence. Proc. Intl. Soc. Mag. Reson. Med. 24. Singapore; 2016:577.

Gurney-Champion OJ, Bruins Slot T, Lens E, van der Horst A, Klaassen R, van Laarhoven HWM, van Hooft JE, Nederveen AJ, Bel A. TU-H-206-08: Quantitative impact of biliary stent artefacts on MR images. Med. Phys. 2016; 43:3775.

Lens E, **Gurney-Champion O**, van der Horst A, Tekelenburg D, van Kesteren Z, Parkes M, van Tienhoven G, Nederveen A, Bel A. MO-FG-BRA-09: Towards an optimal breath-holding procedure for radiotherapy: Differences in organ motion during inhalation and exhalation breath-holds. Med. Phys. 2016; 43:3711.

Curriculum Vitae

Oliver Jacob Gurney-Champion was born on October 31st, 1987, in London, United Kingdom. At age 3, Oliver and his family moved to Hoorn, the Netherlands, where he was raised. He graduated from the pre-university science track of secondary education at the Oscar Romero in 2006. In 2009 Oliver obtained a Bachelor's degree (cum laude) in Physics and Astronomy at the VU University Amsterdam. In 2012 he obtained a Master's degree in Astrophysics at the University of Amsterdam. During his Master's education, Oliver worked on the ionizing effect of black holes. In the same year, Oliver also obtained a Master's degree (cum laude) in Physics (track: Physics of life and health) at the VU University Amsterdam. As part of the Physics degree, he explored a technique for tracking patient's spine motion during radiotherapy. At the end of 2012, Oliver started his Ph.D. research at the Department of Radiology and the Department of Radiation Oncology, both at the Academic Medical Center in Amsterdam. In four years' time, he developed, optimized and implemented several magnetic resonance imaging techniques for radiotherapy purposes, in particular for radiotherapy of pancreatic cancer. His work resulted in this thesis entitled "Magnetic resonance imaging of pancreatic cancer for radiotherapy". From January 2017 onwards, Oliver will be working as Post Doc at the Institute for Cancer Research in London, United Kingdom.

Dankwoord

Vier jaar vlogen voorbij en van de 961 dagen werken, heb ik er niet eentje met tegenzin gehad. Dat komt deels doordat promoveren natuurlijk erg leuk kan zijn, maar het komt vooral door de vele mensen waarmee ik heb mogen samenwerken in de afgelopen vier jaar. Deze mensen zou ik natuurlijk allenmaal uitgebreid willen bedanken. Maar om de druk kosten te drukken bedank ik hier de belangrijkste in het kort.

Om te beginnen bedank ik alle patiënten en gezonde vrijwilligers die voor mijn onderzoeken proefpersoon zijn geweest. Zonder hun was een groot deel van dit werk nooit mogelijk geweest.

Aart en Arjan, bedankt voor jullie uitmuntende begeleiding de afgelopen 4 jaar. Jullie hebben beide de ideale balans gevonden tussen mij begeleiden en mij zelf mijn weg laten vinden. Aart, ik bewonder jouw medeleven in al je werknemers. Bij jou, in het bijzonder, was ik op eerste plaats mens, en pas op tweede plek een werknemer. Ik ben overtuigd dat ik daardoor juist een veel gemotiveerdere werknemer ben. Arjan, van jou heb ik geleerd om doelgericht te werk te gaan en niet te afgeleid te raken in bijzaken. Waar nodig heb je mijn tijd bewaakt. Zonder jouw begeleiding geen 9 hoofdstukken in dit boekje.

Jaap en Coen, bedankt voor de mogelijkheid om te promoveren en het vertrouwen dat jullie in mij hebben gesteld tijdens mijn promotie traject.

Collega's van de afdeling Radiologie (en aanhang van Z0), bedankt voor jullie bijdrage aan mijn werk en voor de erg gezellige lunches, borrels en congressen. Bedankt Abdallah, André, Anne Marije, Anouk, Bram, Claudia, Dennis, Dorien, Eline, Esther, Eva, Geor, Gustav, Henk-Jan, Hyke, Irene, Jasper, Jithsa, Joena, Jordi, Jos, JP, Jules, Jurgen, Kerry, Kevin, Lena, Lucas, Luuk, Marc, Martijn, Martin, Matthan, Michelle, Myrle, Ot, Paul (G), Paul (dH), Pim, Raschel, Sandra, Sofieke, Tanja, Valentina en Wouter. In het bijzonder wil ik Jos bedanken voor al zijn hulp met Mathematica. Bram, Wouter en Dennis bedankt voor jullie geduld, MRI-kennis, pulse programming en Matlab scripts waar ik, zeker aan het begin van mijn project, veel aan heb gehad.

Ondanks dat ik geen bureau had op de Radiotherapie afdeling, was ik daar altijd welkom. Bedankt Anna, Anette, Astrid, Caspar, Daniel, Eva, Edmond, Eelco, Geertjan, Gerben, Isabell, Jan, Judith, Kleopatra, Laura, Lotte, Mélanie, Niek, Peng, Pouya, Silvia, Sophie, Stijn, Tanja en Zdenko, voor alles jullie aan mij hebben geleerd, zowel als de gezellige borrels en congressen/retreats. Astrid, vaak was jij een soort 3^e dagelijkse begeleidster die een probleem vaak weer van een andere kant bekeek, bedankt daarvoor. Geertjan en Eva, bedankt voor jullie bijdrage en artsen blik bij de inteken studies.

Een deel van dit werk was niet gebeurd zonder de studenten, daarom hartelijk dank voor jullie bijdrage J..p.r (ofwel Jasper na een compressed sensing reconstructie),

cijferloze Thabiso en stent-student Thijs, en de 4D-studenten waar ik veel bij betrokken was: Daniel en Isabell.

Sonia, thanks for your visits and our discussions on ATR-SSFP. I really enjoyed working with you.

René, Paul en Jasper, ondanks dat ons werk niet in dit boekje is opgenomen wil ik jullie graag bedanken. De samenwerking met jullie verliep vlot en we hebben toch maar mooi twee nier artikelen weten te produceren.

Hanneke, Remy en Anne, ik mag van geluk spreken dat jullie op het juiste moment zijn komen opdagen op het AMC en interesse hebben getoond in de alvleesklier. Zonder jullie metingen en klinische input was deze thesis lang niet zo uitgebreid geweest. Daarbij heb ik ook heel veel van jullie geleerd. Ik weet zeker dat ondanks mijn vertrek op het AMC, onze wegen elkaar hier niet zullen scheiden. Remy, ik heb genoten van onze scan sessies en je droge humor.

Bedankt Zdenko (Elayne), Lisa (Nynaeve), Jim (Egwene), Laura (Perrin) en Anne-Mieke (Mat). Ik had niet verwacht om op een ESTRO-meeting een 'spelletjes groep' (ik zal DnD hier maar niet noemen, want ik heb het vermoeden dat collega's dit gaan lezen) te vormen. Ik ben blij dat jullie naast collega's (Zdenko, Jim, Laura) en coauteurs (Zdenko en Laura) ook mijn vrienden zijn geworden.

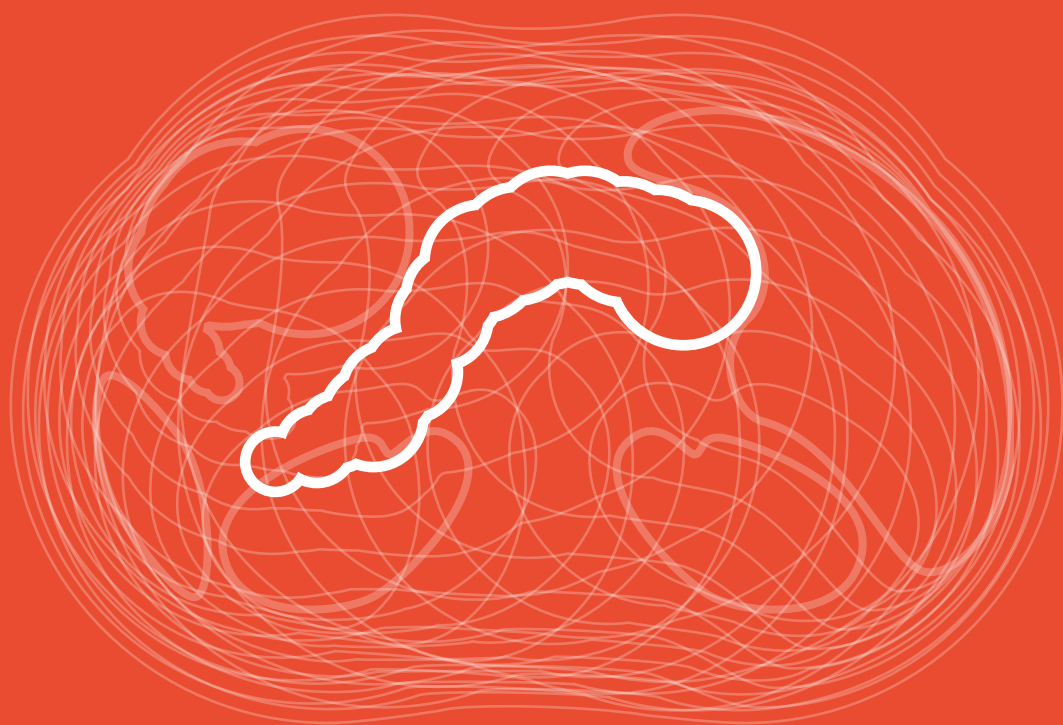
Tom, bedankt voor alle werk en tijd die je in het design van dit proefschrift hebt gestoken.

Lieve familie: moeder/Carolien, paps/Chris en Marjo, broertje/Tom/Tommie/Thomas en Esmay, en opa en oma. Inhoudelijk heb ik natuurlijk niet zo veel aan jullie gehad, maar voor alles om mijn promotie heen had ik geen betere familie kunnen wensen. Ondanks dat ik als kind niet altijd even makkelijk in omgang was en mijn talenten niet altijd even zichtbaar waren voor docenten, hebben jullie altijd mijn sterke kanten gezien en gestimuleerd. Mede door jullie vertrouwen en steun ga ik straks als doctor door het leven. Bedankt voor jullie hulp, bewondering en interesse in mijn werk.

Ik had de afgelopen 4 jaar nooit kunnen volhouden zonder de ontspan mogelijkheden die mijn vrienden mij hebben geboden. In het bijzonder wil ik Marijn, Dirk, Maarten, Mark en Pieter bedanken. Van gamen tot bordspelletjes, van sterke drank tot maltjes, van ballet tot hardlopen, van Duinrell tot Castlefest, van USA tot Wedderveer, van basisschool tot nu, bedankt voor al onze avonturen die altijd gepaard gaan met een goede dosis humor.

Eelco, Caspar en Kelly, mijn medestudenten en kameraden. Wat vond ik het fijn om met jullie te studeren en samen de wereld van studeren, natuurkunde, medische fysica en de wetenschap te ontdekken. Met jullie heb ik gemerkt van 1+1+1+1 meer dan 4 kan zijn, want zonder elkaars enthousiasme en drijf om alles te weten te komen had ik (hadden wij?) het nooit zo ver geschopt.

Lieve Anne-Mieke, wat mag ik van geluk spreken met jou, ook als het op promoveren aan komt. Ik heb altijd mijn werk kunnen bespreken met je, waar je graag op niveau over mee discussieerde. Ondanks dat ik regelmatig tot laat heb gewerkt en vaak later thuiskwam dan we initieel hadden afgesproken, heb ik je hier nooit over horen klagen. Sterker nog, vaak bracht je het als een voordeel omdat je zelf dan ook wat kon werken. In de laatste weken heb je mijn gehele thesis van A tot Z doorgenomen en nog vele stijl/spel fouten weten aan te kaarten, die door velen over het hoofd gezien waren. Vele malen dank voor alles!



ISBN 978-94-028-0546-8