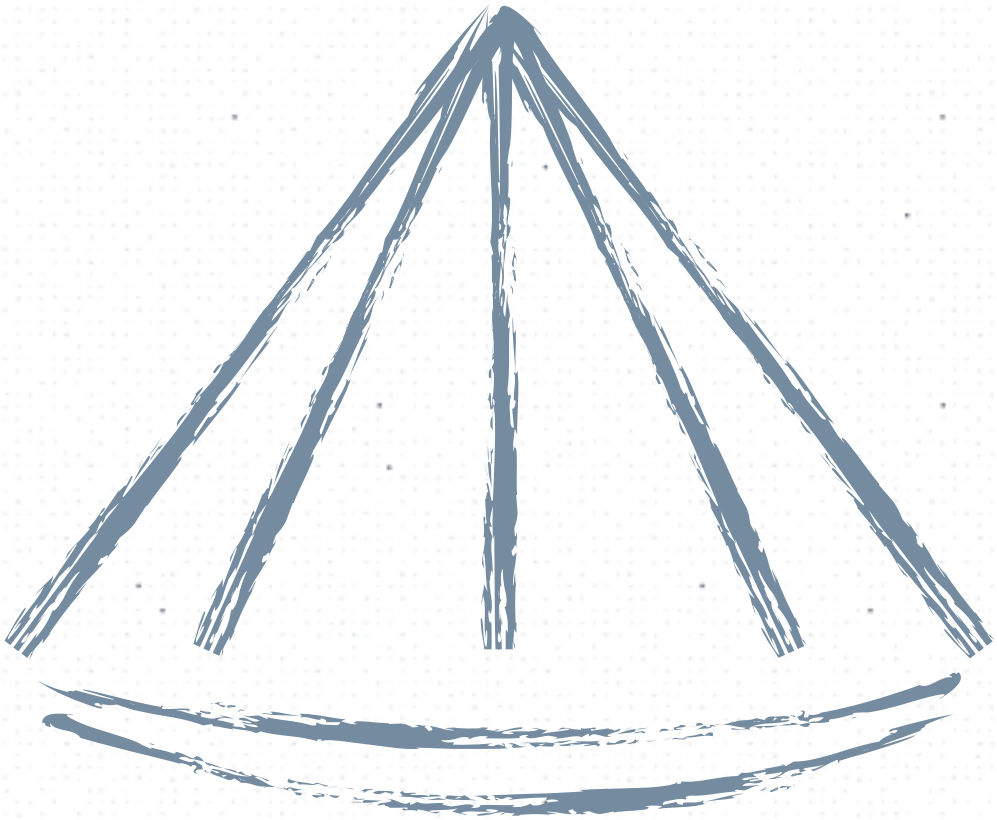


IMPROVING FOCUS ON UTERINE FIBROIDS:

visualizing the treatment effect after ultrasound ablation by using diffusion weighted imaging

Master's Thesis
Technical Medicine
October, 2020

isala

UNIVERSITY OF TWENTE.

D.J. Slotman

Graduation Committee

prof. dr. ir. B. ten Haken	Chair & Technical Supervisor
dr. M.F. Boomsma	Medical Supervisor
ir. J.A.C. van Osch	Technical Supervisor
drs. E.M. Walter	Process Supervisor
dr. A.T.M. Bellos-Grob	External Member

UNIVERSITY OF TWENTE

Abstract

Radiology Department
Isala Zwolle

Master of Science

Improving focus on uterine fibroids: visualizing the treatment effect after ultrasound ablation by using diffusion weighted imaging.

by D.J. SLOTMAN

Uterine fibroids (UFs) are common benign neoplasms of the uterus, entailing a high socioeconomic burden. MR-HIFU is a recent alternative treatment option for symptomatic UFs and has an appealing non-invasive character. Shorter procedure times and improved treatment outcomes could increase the availability of MR-HIFU for patients suffering from UFs. As for now, no established method is available to assess the treatment progress intraprocedurally, leading occasionally to unneeded prolongations of ablations and limited treatment outcomes.

Currently, the treatment effect is examined postprocedurally by acquiring a *contrast-enhanced* (CE)-T1w scan, visualizing the *non-perfused volume* (NPV). A major drawback of a contrast agent is the inability to use it between sonications, due to risk of *gadolinium* (Gd) entrapment and distortion of the thermometry. However, there has been some suggestions about applying *diffusion weighted imaging* (DWI) to visualize tissue perfusion between MR-HIFU sonications. Until now, previous work about DWI has only been analyzed with a mono-exponential model in the context of Gd-free NPV visualization. This thesis has examined the way in which *intravoxel incoherent motion analysis* (IVIM) and *deep learing* (DL) can aid in predicting perfusion based on DWI data after MR-HIFU treatment.

A dataset of 56 patients with a paired DWI and CE-T1w scan was used to conduct this analysis. With the first strategy, a bi-exponential IVIM model was applied. Quantitative analysis disclosed a significant difference between parameters within the perfused and non-perfused volumes, implicating contrast differences between the ablated and viable tissue.

With the second approach, a DL-based method was used to create synthetic CE-T1w scans with a conditional generative adverserial network. With quantitative and qualitative evaluation it was established that the DL-method could produce synthetic CE-T1w scans that can be used for adequate treatment assessment.

These findings show that these methods for DWI-data analysis are feasible for MR-HIFU treatment evaluation of UF. Further work should focus on prospective assessment of these methods and their clinical value. The findings from this work are a next step towards more efficient healthcare by improving the efficiency of MR-HIFU treatments of UFs.

Acknowledgements

Twelve months ago, I returned to the Isala Hospital in Zwolle to join an enthusiastic research team at the Radiology Department to work on further improvements of MR-HIFU. Now, the dissertation "Improving focus on uterine fibroids: visualizing the treatment effect after ultrasound ablation by using diffusion weighted imaging" lies before you, and has been written in attempt to fulfill the graduation requirements of Medical Imaging and Interventions Master's program of Technical Medicine at the University of Twente.

Special thanks goes to my supervisors Martijn Boomsma and Jochen van Osch at the Isala Hospital for their support and excellent guidance during this process. The opportunities I received in the past year are greatly appreciated, and that I was allowed to work on some major projects. Martijn Boomsma showed faith in me by the new role of PhD-student, which I admire. In addition, I wish to thank Elyse Walter and Bennie ten Haken at the University of Twente for their great supervising at the University of Twente at this project.

Furthermore, I would like to thank Erik Phernambucq for his clinical mentoring at the Radiotherapy department in Zwolle, and Wilbert Bartels and Martijn Froeling from the University Medical Center Utrecht for their valuable academic input and support throughout this year.

Finally, I wish to thank my family, friends and Alice in special for being helpful and supportive during my time studying Technical Medicine at the University of Twente.

The accomplishments during this internship make me proud, and I look forward to further contribute to more efficient healthcare around the topics covered in this thesis with the role of PhD-student.
I hope you enjoy reading this work,

Jorik Slotman

Zwolle, August 6th, 2020

Contents

List of Abbreviations

UF	**U**terine **F**ibroid
MR-HIFU	**M**agnetic **R**esonance-guided **H**igh **I**ntensity **F**ocused **U**ltrasound
CE	**C**ontrast **E**nhanced
Gd	**Gd**olinium
UAE	**U**terine **A**rtery **E**mbolization
DWI	**D**iffusion **W**eighted **I**maging
IVIM	**I**ntra**V**oxel **I**ncoherent **M**otion
ADC	**A**pparent **D**iffusion **C**oefficient
MR	**M**agnetic **R**esonance
MRI	**M**agnetic **R**esonance **I**maging
MED12	**MED**diator of RNA polymerase II transcription, subunit **12**
FIGO	**F**édération **I**nternationale de **G**ynécologie et d'**O**bstétrique
US	**U**ltra**s**ound
FV	**F**ibroid **V**olume
PV	**P**erfused **V**olume
NPV	**N**on-**P**erfused **V**olume
ML	**M**achine **L**earning
DL	**D**eep **L**earning
CNN	**C**onvolutional **N**eural **N**etwork
cGAN	**c**onditional **G**enerative **A**dverserial **N**etwork
ReLU	**Re**ctified **L**inear **U**nit
MAE	**M**ean **A**bsolute **E**rror
MSE	**M**ean **S**quared **E**rror
SSIM	**S**tructural **SIM**ilarity
SNR	**S**ignal-to-**N**oise **R**atio
PRF	**P**roton **R**esonance **F**requency
TE	**E**cho **T**ime
PGSE	**P**ulsed **G**radient **S**pin **E**cho
T1w	**T**1 **W**eighted
T2w	**T**2 **W**eighted
ROI	**R**egion **O**f **I**nterest
CI	**C**onfidence **I**nterval

Chapter 1

Uterine fibroids

1.1 Description

A *uterine fibroid* (UF), or uterine leiomyoma, is a common benign fibromuscular tumor of the myometrium. The relatively high prevalence of UF combined with the often occurrence of symptoms requiring treatment induces high socioeconomic costs. Major complaints are abnormal uterine bleeding, pain and discomfort. Both invasive and non-invasive treatment options are available, including drug therapy, surgery, embolization and *high intensity focused ultrasound* (HIFU).

1.2 Epidemiology

UF is the most common type of pelvic tumors in women [1]. Although UF is commonly found, the true frequency is unclear. In literature, a wide range (between 4.5% and 77%) of prevalences of UF is reported, mainly caused by study populations and diagnostic methods [2, 3]. In Europe, the prevalence of diagnosed UF is estimated between 11.7% and 23.6% [4].

In a study processing 100 consecutive hysterectomy specimens with a 2-mm sectioning interval, 649 UFs were found in 77 uteri [3]. Although small UFs may be of no clinical relevance, this study suggests that prevalence measured by other methods may be an underestimation of true UF frequency. It has been estimated that the percentage of asymptomatic UFs excess 50% [5]. Risk of malignancy (sarcomas) for UFs is low: retrospective studies have cumulatively shown 7 cases of unexpected leiomyosarcoma and endometrial stromal sarcoma in 4007 cases of hysterectomy or myomectomy where a benign UF was expected [6].

1.2.1 Socioeconomic implications

Due to the high prevalence numbers of UF in women and the severity of symptoms it regularly causes, the socioeconomic burden is relatively high. In five European countries, research has disclosed a reduction of overall work production of 36.1% and an impairment of general activity by 37.9% [4]. In The Netherlands, the mean costs of hysterectomy following UF diagnosis was €15,779 with indirect costs included (i.e. absence of work), retrieved from the EMMY trial conducted by twenty-eight dutch hospitals between 2002 and 2004 [7]. In the United States, the annual costs of the UF are estimated between 5.9-34.4 billion dollars, including direct, indirect and obstetric outcomes [8]. In a sample of women with symptomatic UF from the United States, women waited on average 3.6 years before seeking treatment, 28% of respondents reported missing work due to their UF symptoms, and 24% thought that UFs prevented them for reaching their career potential [9].

1.3 Etiology

1.3.1 Risk factors

Several risk factors are associated with the development of UF, including age, ethnicity, hormonal factors, weight and lifestyle [1, 10, 11].

Age

In premenopausal women the cumulative incidence of UF increases with age, but this rate decreases at older ages after menopause. This suggest that the premenopausal uterus is less sensitive for developing UF, or that women after an age of 40 years are in the low risk group [12]. A study conducting first-trimester or postmiscarriage ultrasound examinations found that UFs are also common in relatively young women, with a prevalence of 10.7% [13]. Younger age at menarche has also been established to impact fibroid risk: a 1-year increase in age at menarche is inversely related to fibroid risk (adjusted risk ratio of 87%) [14].

Ethnicity

Studies have demonstrated a 2-3-fold incidence and an increased lifetime risk (80% vs 70%) in black compared to Hispanic, Asian and white women. Also, black women have larger tumors at diagnosis, more severe symptoms and earlier age at diagnosis than Hispanic, Asian or white women. [10, 11, 13, 15]

Hormonal factors

The influence of hormones on growth of UF is well-established, in particular the female sex steroids estrogen and progesterone are correlated to UF development [16–18]. It is postulated that factors correlated the overall lifetime exposure to estrogen in woman are risk factors for UF growing [19], such as obesity and early age at menarch [14, 20]. Other lifestyle factors that decrease the total lifetime exposure to estrogen seems to lower risk for development of UF [19]. This includes smoking, although the found inverse association between smoking and UF is weak [21, 22].

Others

Family history, weight, and pregnancy are other factors correlated with development of UF [11]. First degree relatives have an 2.5 times increased risk of UF [23]. The exact relation between weight and UF is shown to be complex, but an inverse J-shaped correlation has been described between UF and BMI in women [24–26]. Literature suggests that pregnancy has an protective effect for UF, hypothetically caused by remodeling of uterus tissue after pregnancy [27, 28].

1.3.2 Molecular and cellular mechanisms of disease

As described in 1.3.1, UFs are largely dependent on estrogen and progesterone hormones. The exact pathogenesis of UF is largely unclear, but the traditional definition of UF is that they are 'clonal smooth muscle cell neoplasms that are growth-responsive to gonadal steroids and have characteristic chromosomal rearrangements underlying their development' [29]. Research continues to expand this view on UF.

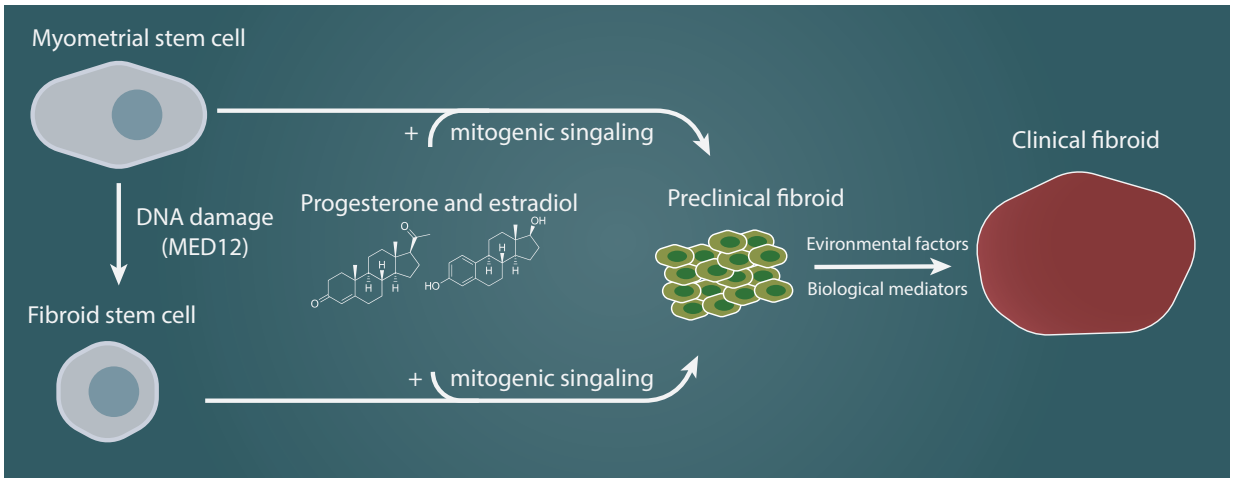


FIGURE 1.1: Schematic overview of UF development. Due to genetic influences, myometrical stem cells can be transformed to fibroid progenitor cells, or fibroid stem cells. Caused by WNT/β-catenin signaling, stem cells in the myometrium cast mitogenic signals under exposure of progesterone and estrogen. From this point, the preclinical UF may eventually develop towards a symptomatic UF, stimulated by environmental and biological factors.

MED12-mutation

The myometrium contains myometrial stem cells that give the uterus its plasticity and regenerative capacity [30]. It is hypothesized that myometrial stem cells are transformed under certain circumstances to fibroid progenitor cells. It is shown that UFs have less stem cells that normal myometrial tissue [31], but contains the somatic *mediator complex subunit 12* (MED12) mutation that contributes to tumorgenesis and genomic instability [32, 33]. The MED12 mutation is shown in 70% of UFs [34]. MED12 is part of the Mediator complex, that forms the bridge between transcription factors and RNA polymerase [35]. UFs with MED12 mutations show higher activations of the WNT/β-catenin pathway compared to normal adjacent myometrium [36]. The WNT/β-catenin pathway is a cell signaling pathway that causes tumor growth of many types [37]. It is demonstrated that WNT/β-catenin signaling has a paracrine role in the growth of UF, by stimulating UF or myometrial cells to send mitogenic signals to adjacent tissue stem cells in response to estrogen and progesterone [38]. Other identified gene groups involved in development of UF are the *high mobility group AT-hook 2* (HMGA2), the *fumarate hydratase* (FH) group and a group associated with deletion of collagen type IV α5 (COL4A5) and COL4A6 [29].

After UF onset with matured UF cells, growth of UFs are stimulated by ovarian steroid hormones. It has been demonstrated that estrogen promotes proliferation of UF through activation of fibroblasts. In these activated fibroblasts, the expression of estrogen receptors were found higher than in smooth muscle cells [39]. Also progesterone plays an important role in development of UFs; it regulates proliferation, apoptosis and collagen deposition of ECM [40, 41]. Besides steroid hormones, growth factors, cytokines and chemokines play an important role as effectors of estrogen and progesterone in initation and development of UFs [41].

1.4 Clinical presentation

1.4.1 Anatomical overview

The nongravid uterus is a pear shaped hollow organ and is located between the urinary bladder anteriorly and the rectum posteriorly with a length commonly around 7 cm [42, 43]. The uterus

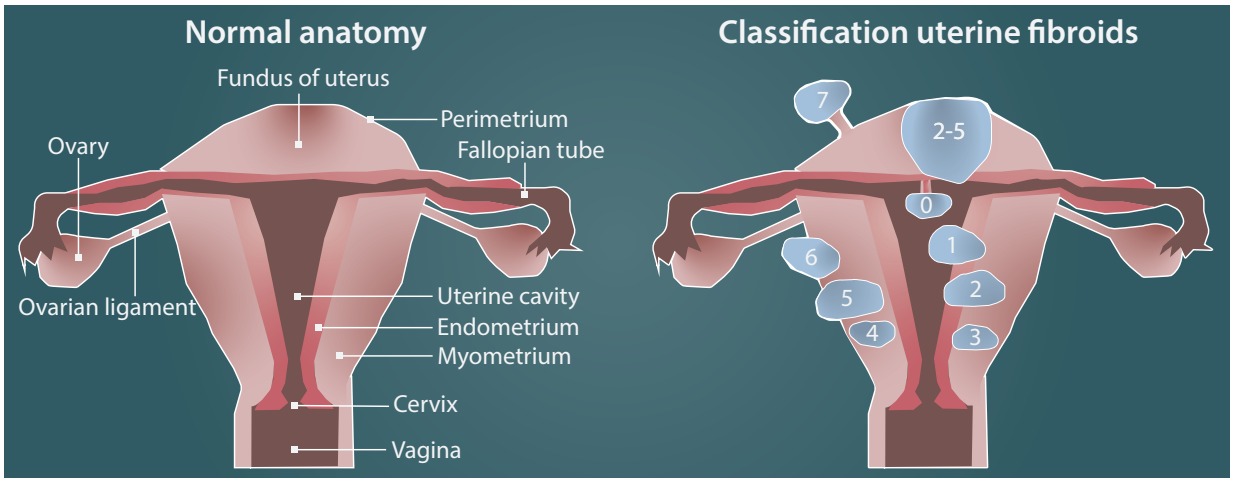


FIGURE 1.2: Schematic anatomical overview of the normal uterus (left), and the uterus bearing UFs (right). For the classification of the UF, see table 1.1

Submucosal	0	Pedunculated intracavitary
	1	$<$ 50% intramural
	2	$\geq$ 50% intramural
Other	3	Contacts endometrium; 100% intramural
	4	Intramural
	5	Subserosal $\geq$ 50%
	6	Subserosal $<$ 50%
	7	Subserosal pedunculated
	8	Other (e.g. cervical, parasitic)
Hybrid	2-5	Submucosal and subserosal, each with less than half of the diameter in the endometrial and peritoneal cavities respectively.

TABLE 1.1: Classification of UFs based on the FIGO classification [45]. See figure 1.2 for the visualization of the different UF types.

can be divided in two main parts: the cervix uteri and corpus uteri, separated by the isthmus. The uterus contains three tissue layers, i.e. the endometrium, myometrium and serosa/perimetrium. The endometrium is the inner lining of the uterus, and responds to the hormonal stimulation and changes during the menstrual cycle. The middle tissue layer is the myometrium, containing the smooth muscle cells UFs arise from. The perimetrium is a very thin outer serosa that is part of the peritoneum. [44] The anatomy of the uterus is schematically drawn in figure 1.2.

1.4.2 Classification

The classification of UFs that is approved by the International Federation of Gynecology and Obstetrics (FIGO) contains two classification systems [45]. The first UF classification system distinguishes submucosal UFs from other locations, because generally submucosal fibroids more often entail bleeding symptoms. The secondary UF classification was originally submitted by Wamsteker et al. [46] and includes 9 different subtypes of UFs. In addition to the distinction between submucosal and other locations, the secondary classification points discriminate intramural, subserosal and transmural lesions (see figure 1.2 in combination with table 1.1).

1.4.3 Symptoms

Most fibroids are asymptomatic. The extent of symptoms of UFs depends on size, location and number of fibroids. Submucosal and intramural fibroids generally cause abnormal uterine bleeding, whereas pedunculated and subserosal fibroids usually entail bulk-related symptoms. Excessive bleeding may lead to anaemia, a potentially life-threatening condition.

Besides abnormal bleeding, (non-cyclic) pelvic pain is also increased in women with UF. There seems no clear relation between number or volume of UF and intensity of pain [47]. Other symptoms besides abnormal bleeding and pelvic pain include dyspareunia, backache, obstructive effects on bladder or rectum and infertility [29].

Acute presentation

UF may cause acute pain, and could be following degeneration of a UF when it outgrows its blood supply, torsion of a pedunculated UF or prolapse of a submucosal UF. Haemorrhagic degeneration can be seen in UF from pregnancy and oral contraceptive pill use due to thrombosis in the venous outflow of the fibroid. In such cases, the UF rapidly increases in volume and acute hemorrhagic infarct occurs. [48] Also intra-abdominal haemorrhage can emerge in rare cases from UFs. It is poorly recognized by clinicians, and may result in hypovolemic shock [49].

1.4.4 Fertility

UFs are found in around 10% of infertile women, and infertile women with no clear cause of infertily have an incidence of 1-2.4% of UFs [19, 50]. The exact impact of UFs on pregnancy is poorly understood, but it is hypothesized that UFs can obstruct the fallopian tubes and inhibit gamete transport or embryo implantation, or distort the endometrial cavity anatomically [50–52]. The ideas that UFs may cause infertility mainly arise from case series, mostly without controls. In these case reports, various previously infertile women conceive after UF removal, and indeed suggest UFs cause infertility, although evidence is weak. [53]

What has been shown is that UFs grow significantly in size during pregnancy, involving 71.4% and 66% between the first and second, and second and third trimester respectively [54]. Moreover, the influence of UFs on fertility depends on the anatomical location of the UF; subserosal fibroids do not affect fertility outcomes, while submucosal UFs decrease fertility. Treatment of submucosal UFs seems to benefit fertility rates (43.3% vs 27.2% for submucosal UFs). [55, 56]

1.4.5 Diagnosis

Diagnosis of UF can be complicated by several factors, including the size, location and number of fibroids. Symptoms of UFs are relatively common and may be caused by a large number of disorders. For example, abnormal uterine bleeding can derive from endometrial and endocervical polyps, adenomyosis, malignancy, coagulopathy, ovulatory dysfunction, primary endometrial disorders and iotrogenic causes [45].

Usually, clinically significant can be found with pelvic examination based on an enlarged, irregularly shaped, firm and non-tender uterus [11], but several imaging modalities are available for differential diagnosis. A strong relationship has been demonstrated between the size estimated based on manual examination and ultrasound, when assessed by an experienced examiner.

Ultrasound

Ultrasound is an often used imaging modality for diagnosis of UF and can be performed transabdominally and transvaginally [57]. Transvaginal ultrasound has generally a higher sensitivity, but transabdominal ultrasound shows a higher sensitivity for fundal UFs [58, 59]. The sensitivity of transvaginal ultrasound depends on observer experience, but is on average around 93%, and the specificity ranges between 73 and 91% [60–62]. Additionally, saline infusion sonohysterography could be a technique that improves the diagnostic accuracy of transvaginal ultrasound. It could be helpful by accurate delineation of the submucosal and intracavitary UFs, and by differentating submucosal UFs from endometrial polyps [59, 63]. With saline infusion sonohysterography, the uterine cavity is filled with normal saline solution (0.9% NaCl) to improve sonographic contrast.

Hysteroscopy

The advantage of hysteroscopy is the ability to directly visualize the uterine cavity and having the option to take biopsies. Disadvantages however are the limited view, since the view depth is limited to the endometrium. However, some general aspects are established that are indicative for a pathological condition, including irregularities, pronounced hypervascularization, fibrous cystic appearance and haemorrhagic cystic lesions. [64] Especially in diagnosis of submucous UFs, hysteroscopy shows excellent diagnostic accuracy, comparable to sonohysterography, with an higher accuracy than conventional transvaginal ultrasound [65].

Biopsies can be also be collected transvaginally and transabdominally by US guidance. Preoperative biopsies can aid differentiation between UFs, adenomyosis and sarcomas. However, especially with this method it should be considered carefully if the invasive biopsy is in balance with the clinical benefit of early diagnosis. [66]

1.4.6 Magnetic Resonance Imaging

Due to accessibility and low costs, US is the first diagnostic imaging modality in line. Nevertheless, *magnetic resonance imaging* (MRI) has better soft tissue contrast, a larger field-of-view and can display multiplanar images in comparison to US, therefore it can help clinicians in pretreatment planning and differential diagnosis of adnexal masses [67, 68]. It has been established that despite US being an efficient tool for UF diagnosis, when UFs are large or multiple, MRI mapping is superior [69].

The guidelines of the European Society of Urogenital Radiology (ESUR) propose an MRI-protocol with at least two T2w orthogonal planes of the uterus including a sagittal sequence of the uterine corpus. This sagittal sequence allows other sequences to align with the axis of the uterus, instead to the body. Besides the T2w sequences, an axial T1w sequence is advised for the assessment of other pelvic pathologies and high-intensity lesions. Also *gadolinium* (Gd) iv can be indicated for UFs with rapid grow, high intense areas on T2w, differentation between an adnexal mass and pre- and post-treatment assessment. [70] In figure 1.3, an example sagittal T2w MRI scan is depicted with two UFs from a 51 year old patient.

Classically, UFs are sharply demarcated on a T2w sequence and have a lower signal intensity compared to the myometrium. In some intramural and subserosal UFs a high intensity rim could be visible, composed of edema and dilated vessels. UFs can become degenerative when they outgrow their blood supply, and can be suggested by an heterogeneous UF appearance on T2w scans. [71–73].

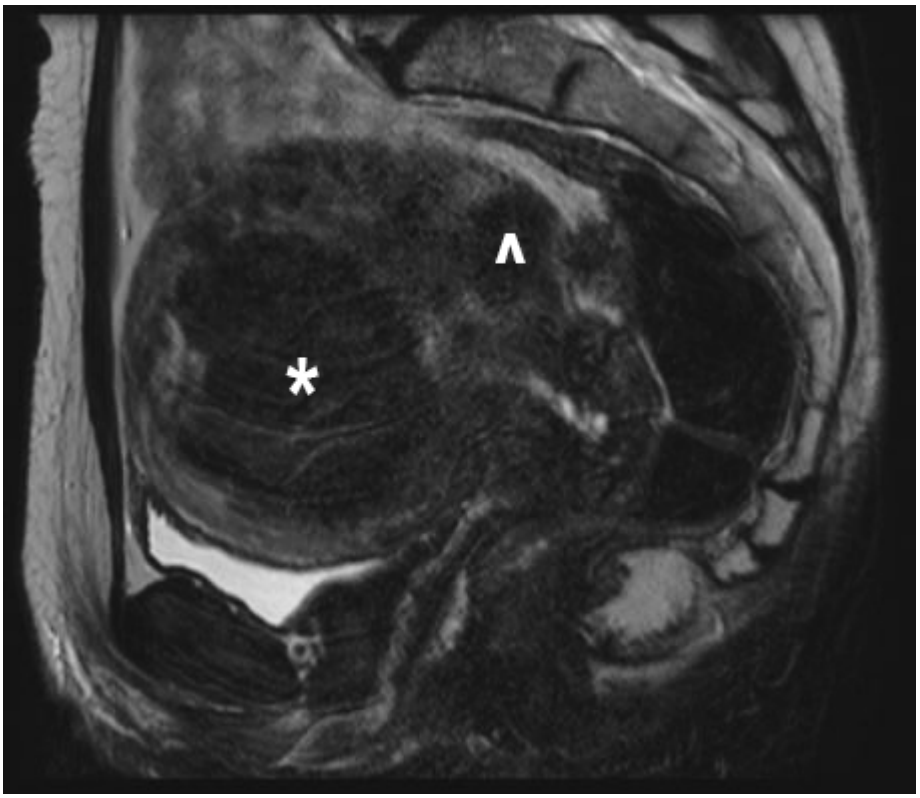


FIGURE 1.3: Sagittal T2w slice of two UFs with both Funaki type 2 in a 51-years old patient. FIGO class 2-5 (*) and 4 (^), and 263 (*) and 14 (^) cm^3.

Unfortunately, differentiation between atypical UF and uterine leiomyosarcoma is difficult on MRI. Four qualitative MR features have been proposed that show strong statistical association with leiomyosarcoma at histopathology (nodular borders, haemorrhage, 'T2 dark' areas and central unenhanced areas. Three or more present features accurately distinguish atypical UF from leiomyosarcoma. [74]

1.5 Treatment of uterine fibroids

Nowadays, multiple treatment options for UF with various degrees of invasiveness have been established. Due to the large scope of the disease, many attempts have been made to find efficient treatment methods during history. [75] It has generally been accepted that patients with asymptomatic UF and no desire for pregnancy only require periodic monitoring and no special treatment, due to the largely benign nature of UFs. Therefore, care should be taken to select the least aggressive options minimizing risks and optimizing outcomes. [76, 77]

1.5.1 Medical treatment

Since 1940 attempts have been made to treat UFs medically with progestins and estrogen-progestins combinations [78], but evidence for their effectiveness is lacking. When discussing estrogen-progestin combinations (oral contraceptives) as treatment for UFs, two questions arise: whether their use can prevent arising of UFs, and whether they can decrease the size of already existing UFs. There is no large body of literature answering positive or negative on either question. Therefore the current concept is that oral contraceptives are safe to use in this regard, but have no inhibitory effect on UFs either. [77, 79]

GnRH agonists

Godanotropin-releasing hormone (GnRH) agonists are structurally similar to the natural GnRH hormones synthesized in the hypothalamus, stimulating excretion of gonadotropins from the anterior pituitary. GnRH agonists have been proposed as conservative treatment option for UFs in 1983. It causes suppression of gonadal activity by desensitization of GnRH receptors after an initial increase in the release of gonadotropins. The desensitization leads to a hypogonadal state that simulates menopause. [80] It has been shown that GnRH agonists shrink UFs significantly, but the the downside of GnRH agonists usage as treatment besides the vasomotor symptoms is the regrowth of UFs and recurrence of symptoms. Additionally, the usage can not exceed 6 months due to the increased risk of osteoporosis due to hypoestrogenism. [81, 82] It has extensively investigated if GnRH can be used as pre-treatment medication, and it has been shown beneficial for several procedures through uterine and fibroid volume reduction. [77, 83]

Another type of medicine in the same category are GnRH antagonists. They cause immediate suppression of gonadotropins and do not lead to the peak of gondatrophins as with GnRH agonists. A disadvantageous fact about the GnRH antagonist as medical treatment for UF is the need to be present in the bloodstream all the time to prevent the action of endogeneous GnRH. [84]

Selective progesterone receptor modulator

Various *selective progesterone receptor modulators* (SPRMs) have been tested for their effect on UF reduction, including mifepristone, ulipristal acetate and asoprisnil. As stated earlier, progesterone plays stimulates growth of UF, in addition to estrogens. SPRMs bind to progesterone receptors and function as progesterone antagonist. SPRMs have been evaluated as pregnancy termination drug, but also as therapeutic option in the treatment of UF. A review of 14 RCTs and 1021 included women with UF concluded that SPRMs improved fibroid-related symptoms and quality of life, and decreased menstrual bleeding in comparison to placebo. No conclusions could be drawn about changes in pelvic pain. [85] Usage of SPRMs could lead to endometrial changes, but these changes reversed when SPRM treatment was discontinued. These endometrial changes consist of mild thickening of the endometrium and cyst, gland and vascular changes, but are not considered dangerous [85, 86].

1.5.2 Surgical interventions

Surgical interventions are still the main treatment strategies for UF, including hysterectomy and laparoscopic of hysteroscopic myomectomy [87].

Hysterectomy

After cesarean section and removal of the fetus or other procedures involving delivery, hysterectomy is the most common major surgery in the United states. UFs are the most common diagnosis with hysterectomy, in around 30-40% of the cases. [88, 89] Hysterectomy is a surgical procedure to remove the uterus and in some cases the cervix. Total hysterectomy refers to the removal of the uterus and (partial) cervix, while a subtotal hysterectomy involves removal of the uterus without the cervix. This treatment option is the most radical and definitive for UFs, and prevents the recurrence of UFs.

Hysterectomy can be performed via the abdominal, vaginal and laparoscopic route, eventually robot-assisted. Vaginal hysterectomy may result in faster procedures, shorter discharge

time and smaller blood loss. However, the size of UF limit the applicability of vaginal hysterectomy. [90] Studies with one-year follow up consistently show that outcomes of hysterectomy are good: low risk of complications and improved quality of life. However, concerns have been raised about the long-term consequences, regarding increased risk of cardiovascular disease and dementia [87, 91]. Fed by these concerns, risk of overuse of hysterectomy as UF treatment is issued, including the notion that removal of the female reproductive organs should be considered carefully [92].

Myomectomy

Myomectomy is a fitting treatment option for patients with submucosal fibroids that wish to preserve their uterus or fertility. With hysteroscopic myomectomy, an endoscope is placed through the cervix to remove the UF(s) by intruding the endometrial cavity. This is performed under visual sight via a camera. [87] This procedure is preferred for UFs with a diameter of 6 cm or less [93], and can be performed abdominally and hysteroscopically. A large part of the UF should be intracavitary (FIGO class 0 or 1). Removal of the submucous UFs appears likely to improve fertility [94]. UF recurrence rates are cumulatively around 12%, 36%, 53% and 84% after one, three, five and eight years after laparoscopic myomectomies. Cumulative proabilities for reoperation for UF are around 7% and 16% after five and eight years. [95]

1.5.3 Uterine artery embolization

Uterine artery embolization (UAE) is a minimally invasive procedure that spares the uterus and involves occlusion of both the uterine arteries with particulate emboli. This results in ischemic necrosis of the UF, but no permanent effect on the uterus. To execute the embolization, an angiography catheter is inserted in the patients femoral artery, and the contralateral uterine artery is embolized. This process is repeated for the ipsilateral site, to the femoral puncture. The most common material for the occlusion agent is polyvinyl alcohol with a size of 150-1000 microns. Blockage of the arteries can be confirmed by angiography. When both uterine arteries are occluded, the normal myometrium recovers blood flow by fast establishment of an alternative blood flow via ovarian or vaginal blood flow, while the UFs are provided by end arteries and cannot re-establish a collateral blood flow. This can lead to ischemic necrosis of the treated UF, and eventually symptom relieve of the patient. [96]

Overall, patients have good procedure satisfaction for UAE, and effective post-operative symptom improvement. Post-operative complication rates are around 18% and 23% for small and large UFs. Reintervention rates are between 2.5 and 8.6% for small and large UFs respectively, and increase to 28% after five years [97, 98]. Studies show tentative indications that UAE has a decreased chance of pregnancy compared to myomectomy [96, 99].

1.5.4 High intensity focused ultrasound

An alternative non-invasive treatment option for UF is *MR-guided high intensity focused ultrasound* (MR-HIFU), and achieves thermal necrosis by delivering local acoustic energy. Short- and long-term results are promising in unrestricted treatments [100], and this treatment option will be discussed in depth in chapter 2.

Chapter 2

MR-HIFU and DWI

2.1 MR-HIFU

Studies over the past century have established the ability of very short focused soundwaves to heat tissue non-invasively. It was discovered that by focusing ultrasound waves with a concave surface, ultrasound energy at the focal spot could be concentrated 150 times as much than at a point close to the source. This discovery from Grutzmacher [101] in 1935 gave rise to the study from Lynn et al. in 1942 [102], who showed biologically that focused ultrasound can be used as method for ablating tissue in living animals. From that time, several clinical studies were conducted, ranging from neurosurgery, ophthalmology to oncology [103–105]. Widespread clinical used was hindered by the parallel development of alternative treatment options [106]. Several technological advances, including MR-guidance and -thermometry, made HIFU an easier and safer approach and a potential treatment option, such as palliative treatment of metastatic bone disease and UFs [107–111].

2.1.1 MR-HIFU treatment of UF

MR-HIFU is a non-invasive, organ-sparing, trans-cutaneous ablation technique. Since the first feasibility study of MR-HIFU treatment of UFs was published in 2003, many MR-HIFU procedures of UF ablation have been performed [112]. MR-HIFU treatment of UF received a European CE mark in 2002 and FDA approval in 2004 [113]. Since then, MR-HIFU have become a major non-surgical treatment option for UFs, and have an evident appeal of complete non-invasiveness [114].

Patients with an (US) confirmed diagnosis of UF in the pre- or peri-menopausal state can receive a MR-HIFU treatment. Not all patients are eligible: an UF with a diameter < 1 cm, current pregnancy, an abdominal fat layer > 4 cm, MRI contra-indications or high T2 signal of the UF (assessed with Funkaki classification) are examples of inhibitors of MR-HIFU treatment.

Funaki classification

It is well-established that the response of UF to MR-HIFU treatment is correlated to the T2 signal intensity [115]. Funaki et al. (2007) [114] proposed a classification system for UF based on T2 signal intensity that is widely adopted into clinical practice. This classification consists of:

- Type 1: low T2 signal, i.e. comparable to skeletal muscle.
- Type 2: intermediate T2 signal, i.e. higher than skeletal muscle, but lower than myometrium.
- Type 3: high T2 signal, i.e. equal or higher than myometrium.

This characterization of UFs aid the response prediction of MR-HIFU. Funaki et al. argued that UFs of type 1 and 2 are suitable candidates for MR-HIFU treatment, while type 3 are not.

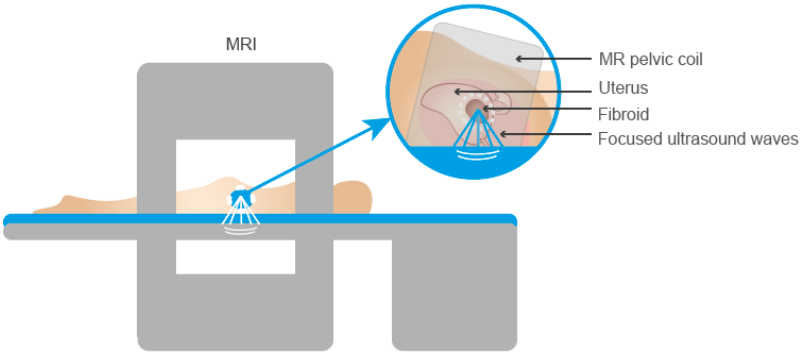


FIGURE 2.1: Schematic illustration of the HIFU-procedure inside an MRI-scanner of a UF. The patient is in prone position, with the US transducer inserted in the HIFU-table. Image adopted from Profound, Sonalleve ®.

Procedure

An example of a MR-HIFU procedure for UF treatment is described in this paragraph. At moment of screening, an MRI is performed to assess the location and UF characteristics. Before treatment, patients fast 6h and are asked to shave their lower abdomen and pubic area when eligible. Patients receive premedication (paracetamol 1000 mg, diclofenac 100 mg and oxycodone 10 mg). A catheter and intravenous line are also inserted. The ablation procedure is executed in an MRI-scanner, as illustrated schematically in figure 2.1. Patients are positioned in a prone position on a HIFU table, that is placed over the MRI table. The HIFU US transducer is inserted in the HIFU table, within a water tank. A membrane lays on top of the transducer, with actively cooled water in between, to reduce the risk of skin burns. Prior to the treatment, conscious sedation (propofol-fentanyl) is administered to the patients during the procedure to control patient's reactions and movement. The patients are asked to lie still, and report any discomfort or pain. At first, a screening T2w scan is made to assess the position of the UF, that may deviate from the position during screening due to the mobility of the uterus. With several manipulation techniques, the location of the UF can be adjusted to a certain degree. A protocol for manipulation is recently proposed by Verpalen et al. (2020) [116].

When the positioning of the UF is finished, the ablations are started. Different ablation cell sizes (focal spot areas) can be chosen, and the the location of the acoustic focus can be electronically controlled. It has been shown that energy efficiency improves with cell size, independent of distance to the US source or T2 signal intensity [117]. After the sonications, when the ablation is expected to be finished, a post-procedural *contrast-enhanced* (CE) T1w scan is acquired to visualize the treatment effect, the *non perfused volume* (NPV). With the NPV, the non-perfused ratio (NPV%) can be calculated by dividing the NPV by the UF volume. Finally, the patients are transferred to a clinical ward, and typically discharged on the same day.

Side effects

Side effects of MR-HIFU treatment are relatively rare, potential side effects include:

- peri-interventional pain (usually low and of short duration), can usually be managed well with analgesia,
- (slight) skin burns,
- slight inflammation of subcutaneous fat tissue and of the abdominal muscles,

- paraesthesia of the legs due to irritation of, or damage to, the nerves,
- deep vein thrombosis of the legs (very rare),
- intestinal lesions or intestinal perforation (extremely rare),
- discharge of vaginal tissue,
- increased and/or irregular bleeding for around 3 months post procedure.

These side effects are adopted from [118]. The most common complication are skin burns, abdominal pain or discomfort, sciatic nerve paresthesia, or leg pain [119].

2.1.2 Treatment outcomes

In general, evidence of treatment outcomes for MR-HIFU treatment of UF is still poor to moderate, mainly because the existing literature has a case series methodology. The point estimate of 18 articles of NPV% is 68.1% (59.9%-76.0%). The mean sonication time of 10/18 studies was 145.6 minutes. Re-intervention percentages after 3-22.6 months post-HIFU ranged from 0 to 21%. All the studies described UF shrinkage, and the *transformed symptom severity scores* (tSSS) where 46.1% after 3 months, 56.1% after 6 months and 53.6% after 12 months. [120] The outcomes of several studies writing on MR-HIFU ablation of UFs must be interpreted with caution, as not all study sites aimed for complete UF ablation, due to earlier regulations. Therefore, distinction should be made between targeting complete or incomplete ablation. When focusing on studies aiming for complete ablation only, it can be concluded that MR-HIFU is a safe and effective method for symptom reduction of symptomatic UF, in a non-invasive manner. Other reviews endorse the proposition that MR-HIFU is a safe, promising and effective therapeutic technique for decreasing of UF volume and symptoms [120]. However, more systematic research is needed to move the debate forward [119, 121].

Fertility

Literature on fertility after MR-HIFU treatment of UF is scarce, but fertility may be reserved after MR-HIFU, studies disclosed [118, 122, 123]. In the study from Rabinovici et al. (2010) [122], 54 pregnancies in 51 women occurred after MR-HIFU treatment of UF. The mean time of conception was 8 months after treatment, life birth was present in 41% of pregnancies with a 28% spontaneous abortion rate, 11% elective pregnancy terminations, and 20% ongoing pregnancies beyond 20 gestational weeks. Another study from Li et al. (2017) [123] described 131 from 189 patients that got pregnant with a 76.3% full-term birth rate and a cesarean section rate of 72.0%. The authors advocate that (nulliparous) women who undergo MR-HIFU treatment of UF can have successful pregnancy and deliver safely.

Cost-effectiveness

Although data regarding cost-effectiveness is limited, a few studies have shown that MR-HIFU treatment of UF can be equally assumed cost-effective in a five year time frame, compared to UAE and myomectomy [124, 125]. The study from Cain-Nielsen et al. (2014) [124] conducted in the United States, MR-HIFU, UAE and myomectomy costed on average respectively $21,232, $22,819 and $22.599.

2.1.3 Physical principles of HIFU

The main goal of HIFU as non-invasive tissue ablating technique, is to selectively destruct tissue and keeping the surrounding tissues intact. This is achieved by using high frequency sound waves, that propagate through a medium. Instead of using catheters or antennas to deliver energy as with other minimally invasive techniques, MR-HIFU uses sound waves to propagate energy through a medium. When sound waves propagate through a medium in the x direction, the intensity I decreases exponentially with distance. This phenomenon is called attenuation, and is described by:

$$I_x = I_0 e^{-\alpha x} , \tag{2.1}$$

where I_0 is the incident intensity, α is the overall attenuation coefficient and x the traveled distance through the medium. The pressure fluctuations of the wave lead to a microscopical shearing motion within the tissue, causing friction and thereby heat. This means that a part of the mechanical energy carried by the soundwave is absorbed by the tissue, constituting the main principle of ultrasound-induced hyperthermia. Other causes of attenuation of soundwave by propagating through a medium are scattering and reflection [126].

The total attenuation α is the sum of attenuation caused by absorption, scattering and reflection. This can be viewed as follows:

$$\alpha = \alpha(absorption) + \alpha(scattering) + \alpha(reflection) , \tag{2.2}$$

Attenuation is strongly dependent on frequency, since absorption and scattering also frequency dependent. This dependence can be summarized with:

$$\alpha = k f^{1.1} , \tag{2.3}$$

where k is a tissue-specific constant, and f the frequency. Note that the intensity decline is often displayed on the logarithmic decibel (dB) scale. The relation between the logarithmic overall attenuation coefficient μ and α is: $\mu = -4.3\alpha$. [127] This frequency dependency makes ultrasound well-suited as mechanism for a non-invasive treatment. The depth of application of the ultrasound energy can be increased by decreasing the frequency of the incident sound waves.

The use of sound waves as method of energy transfer has been established as safe within the diagnostic frequencies [128, 129], and clinically for MR-HIFU a beam around 1.5 MHz is typically used. *Sonic intensity* is a measure for the time-average rate of sonic energy-flow through an area, and expressed in W/cm^2. The sonic intensity of diagnostic ultrasound is often lower than 0.1 W/cm^2 [130], while MR-HIFU can deliver 10^3 to $10^4 W/cm^2$ at the focal spot [131].

The goal of MR-HIFU is to maximize the energy delivery at the focal spot, while minimizing the heat deposition in surrounding tissues. Ultrasound waves interact with tissue it propagates through by the particle motion and pressure variation. Since the ultrasound waves are continuously losing energy through absorption, the temperature of the medium elevates. When this temperature rise is above a certain threshold and maintained for an adequate period, the tissue damages. This thermal effect can be used for coagulation or ablation, and is similar to other ablative techniques that deliver energy to tissue. The threshold at which thermal damage occurs depends on tissue type (absorption and attenuation coefficient) and physiological factors such as pH and O_2. Moreover, the heat build up is highly depending on local blood perfusion rate. At short ablation times, the influence of blood perfusion is small and the heat transfer is mainly causes by thermal conduction. When exposure time increases, the influence of blood perfusion is increased, and this relation is visualized in figure 2.2. [132]

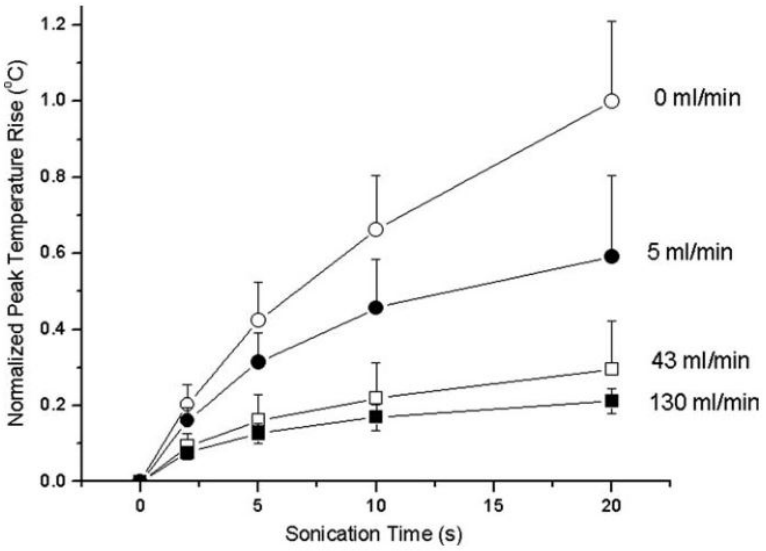


FIGURE 2.2: Relative temperature rise in relation to sonication time in seconds, measured at 11 locations in 5 kidneys for different perfusion flows. Adopted from [133]

2.1.4 MR thermometry

During the application of a thermal dose, monitoring the temperature distribution is paramount. This is possible with MR thermometry, based on the proton resonance frequency (PRF) shift of the water proton. The PRF has an almost linear dependency on temperature, regardless of the tissue [134]. Due to an increasing temperature, Brownian motion in increased, and the bonds of hydrogen are stretched, broken and bent, resulting in a shift in PRF. Caused by the changing bonds, the average time the water molecules are in a hydrogen-bonded state is reduced, and thus decreasing the local magnetic field and thereby the PRF of water. [135]

MR-phase images help deriving the temperature change, by finding the difference between the successive phase images:

$$\Delta T = \frac{\Delta\Phi}{\alpha\gamma T_E B_0}\,, \tag{2.4}$$

with Φ being the change in PRF, α the temperature dependent water resonance chemical shift (0.0094 ppm/°C), γ the gyromagnetic ratio ($42.58 MHz/T$), T_E the echo time and B_0 the magnetic field strength. Since the temperature is measured relative to another time, the tissue needs cooling down to body temperature to prevent unforeseen heat build up, with possible skin burns as a result.

Besides PRF, other quantitative temperature measurements can be performed with MR-techniques, including techniques based on altered Molecular Diffusion Coefficient:

$$D \approx D_0 e^{\frac{E_a}{kT}}\,, \tag{2.5}$$

where D is the diffusion coefficient, D_0 the initial diffusion value, E_a the activation energy of the material, k the Boltzmann's constant and T the absolute temperature in Kelvin. The coefficient D can subsequently be derived by *diffusion weighted imaging* (DWI) with MRI, see section 2.2.

2.1.5 Limitations HIFU

Limitations of MR-HIFU treatments are mainly caused by contra-indications. First of all, UFs with relatively high T2 signal (Funaki type 3) are often excluded from MR-HIFU treatment, since adequate heat build-up is inhibited. Also, application of MR-HIFU treatment is limited by UF size, mainly because of the risk of deep venous thrombosis. Although this risk is relatively small, a large number of UFs or UFs with a diameter larger than 10 cm are controversial for MR-HIFU treatment [136, 137]. Another reason for limitation of MR-HIFU application is the technical aspect that the MR-HIFU beam typically cannot reach deeper than 12 cm from the skin. This also dictates the incapability of MR-HIFU to treat UFs from patients with a large abdominal subcutaneous fat layer. Therefore, UFs that are distant from the skin, or too close to the sacral bone cannot be treated. [138]. Earlier, bowel interposition was also a limiting factor, but new methods have proposed methods to minimize this problem [116]. Other factors limiting the use of MR-HFU are abdominal scarring, concomitant adenomyosis, a postmenopausal state and pedunculated UFs.

Targeted vessel occlusion

Methods are searched to decrease the overall duration of the MR-HIFU treatment, to make this treatment option more cost-effective, compared to the alternatives [139]. There are suggestions that so called targeted vessel occlusion might increase the efficiency of MR-HIFU treatment of UF. With this proposed method, MR-HIFU is intentionally aimed at the uterine arteries supplying the UF. Preliminary results indicate that this might be a feasible and eventually cost-effective to maximize the devascularization ratio of UF within acceptable treatment times. [140]

However, one limitation to this approach is the current inability to clearly visualize the NPV during treatment. This is entailed by the fact that currently a Gd-based contrast agent is needed for clear visualization of the devascularized region of the treated area. However, due to safety concerns (i.e. trapping, dissociation and long-term deposition of Gd), it is often adopted that MR-HIFU treatment cannot be continued after injection of a Gd based contrast agent. Although some argue that MR-HIFU can be performed carefully after administration of Gd, thermometry cannot be adequately performed after Gd injection due to susceptibility artifacts caused by Gd [141, 142].

2.2 Diffusion weighted imaging (DWI)

2.2.1 Diffusion

Around 60%-70% of the human body exists of water. In the complex environment of the human body, water molecules are divided between cells and extracellular compartments. The water molecules move, either in an 'orderly' manner, of randomly.

Random motion

Due to the inherent thermal energy, water molecules at body temperature move constantly. When no obstacles are present, they would move infinitely in the same direction (Newton's first law of motion). But in a medium, water molecules collide constantly and change in direction. This occurs in a short space of time and very often due to the amount of molecules present, and is therefore too complex to predict exactly. From a practical standpoint, these events can be viewed effectively as random motion, and is therefore sometimes called *random walk*. The

random motion of particles is referred to as Brownian motion. This Brownian motion is independently, and has no preferred direction.

The macroscopic manifestation of Brownian motion is called diffusion, and describes the movement of particles as a result of Brownian motion. When no obstructions are present, the diffusion type is called *free diffusion*, or *isotropic diffusion*, the particles can move in each direction with only other water molecules obstructing them. At free diffusion, the distribution of molecules has a Gaussian shape with an average of zero. This means that water molecules are less likely to travel great distances than small, and the largest chance is to have a net displacement of zero.

Restricted and hindered diffusion

When obstacles are present in a medium with impermeable boundaries, such as inside a cell membrane, the diffusion is restricted. When these water molecules are trapped inside a cell membrane for example, the motion of water is determined by the geometry of the cell membrane.

Another type of restriction is called the hindered diffusion. This occurs when the diffusion is hindered by objects, but not completely confined. For instance, this occurs in the extracellular space, or the interstitium. In this case, water molecules can diffuse in all directions, but are impeded. That means that eventually the water molecules can reach any distance, but the progress can be slowed down in certain directions. This impeded form of diffusion depends on the shape of the boundaries.

From the first law of Fick, the diffusion flux J in $kg/(cm^2 s)$ is given by:

$$J = -D\frac{\Delta C}{\Delta x}\,, \tag{2.6}$$

where C is the concentration and x is the position. Thus, the diffusion flux describes the net rate of particles in kg moving through a certain area, and the gradient describes the particles in kg per volume per distance x. This means:

$$[\frac{kg}{cm^2 s}] = [-D][\frac{kg}{cm^4}]\,, \tag{2.7}$$

and implicates that:

$$[\frac{1}{s}] = [-D][\frac{1}{cm^2}]\,, \tag{2.8}$$

and the units of the diffusion coefficient are:

$$[D] = [cm^2/s]\,. \tag{2.9}$$

Diffusion forms the basis of *diffusion weighted imaging* (DWI), and bears often important clinical information [143].

2.2.2 MR-imaging of diffusion

With MRI, imaging relies on interaction with hydrogen nuclei inside a volume. (Sub)atomic particles contain a quantum mechanical property that is called 'spin'. Spin results in a magnetic moment for atoms with an odd mass number (including hydrogen H^1) a nuclear moment.

When placed inside a magnetic field, these spin properties result in a precessing motion with a frequency proportional to the magnetic field strength. This is called the Larmor precession, and forms the basis of the magnetic resonance imaging, including DWI:

$$f = \gamma B \,, \tag{2.10}$$

where f is the frequency in Hertz, γ is the gyromagnetic ratio in Hertz/Tesla and B is the field strength in Tesla. When precessing particles are inside a (strong) magnetic field, they together result in a net magnetization. When a gradient is applied across a sample inside this magnetic field, the net magnetization changes in recession frequency, as can be easily seen from equation 2.10. This principle is used for DWI.

Pulsed Gradient Spin Echo Sequence

The *Pulsed Gradient Spin Echo* (PGSE) is often used for applying a diffusion weighing to an MRI signal. With PGSE, a pulsatile magnetic gradient is added before and after the refocusing pulse of a spin echo, as visible in figure 2.3

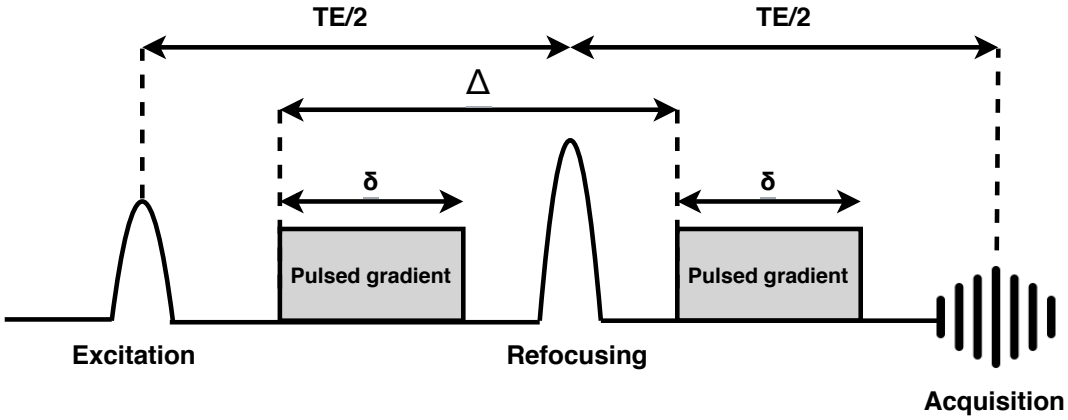


FIGURE 2.3: Schematic illustration of the Pulsed Gradient Spin Echo, often used for DWI imaging. A gradient is added before and after the refocusing pulse, adding a phase difference within the sample. When no diffusion is present, both the gradient pulses should cancel each other out, resulting in no signal attenuation due to diffusion. If diffusion effects are present, MRI signal is attenuated at the moment of acquisition, since the dephasing is not completely canceled out as particles have changed location between the first and second gradient pulse.

After the excitation pulse, the net magnetization is rotated onto the transverse plane. Thereafter, the first pulsed gradient is added that dephases the spin magnetization, due to the different field strengths along the gradient. After the pulsed diffusion-encoding gradient, the refocusing pulse is applied, in order to recover the transverse magnetization. Then, the second pulsed diffusion-encoding gradient is applied and lastly the echoed signal is read out.

When no diffusion is present, magnetization is completely rephased by the two pulsed gradients, since the molecules have not changed location and did experience the same gradient field. However, when diffusion is present in the sample, the molecules will move between the first and second gradient pulse. Hence, the experienced gradient field has changed and the frequency of precession will be adapted (equation 2.10), resulting in a phase difference before the first and after the second gradient pulse. This causes an attenuated transversal magnetization

during readout. Differences in net magnetization attenuation are correlated to differences in diffusion of the measured sample. [144]

b-value

To quantify the influence of diffusion in a MRI sequence, the b-value forms the standard single measure. The b-value describes the weighing of diffusion inside the MRI-signal, or the sensitivity to diffusion. Low b-values have low signal attenuation due to diffusion compared to high b-values. This means that the amount of signal attenuation is increased with an increased b-value, when diffusion is present in the sample. A b-value of 0 s/mm^2 will have no attenuation caused by diffusion in the MR signal. On the opposite, a b-value of 1000 s/mm^2 are usually required to measure slowly moving water molecules [145]. Appropriately, low b-values contain a large portion of the signal in relation to the signal when no gradient pulse is applied, in opposite to high b-values. This entails a relatively low SNR for low b-values.

The signal equation of PGSE was derived by Stejskal and Tanner (1965) [146], for liquids with unrestricted diffusion:

$$I = I_0 e^{-bD} , \tag{2.11}$$

where I is the signal intensity, S_0 the initial signal intensity, b the b-value and D the diffusion constant. The b-value is defined as:

$$b = \gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3}) , \tag{2.12}$$

where δ is the gradient duration, G the gradient strength, δ the gradient duration and Δ the time between the start of the first and second pulsed gradient, as depicted in figure 2.3. Thus, in the b-value, the three parameters G, δ and Δ are incorporated. With these parameters, the amount of diffusion weighted can be controlled of the PGSE sequence. The contrast of the DWI-scans can be changed with these parameters, and can be optimized for specific applications. [144]

2.2.3 Apparent diffusion coefficient (ADC)

When diffusion is not freely, the diffusion coefficient D deviates from the standard value (around $3.0 * 10^{-3} cm^2/s$) at 37 °C [147]. In biological tissue, the water molecules can not move freely due to obstacles, and the actual diffusion distance is limited and slowed down. As explained in the above paragraphs, the diffusion can be measured with MRI via signal attenuation. The signal attenuates highly if fast diffusion is present, and conversely attenuation is small when slow or no diffusion is present.

Accordingly, with the signal attenuation acquired with a diffusion weighted MR-sequence (e.g. PGSE) the *apparent diffusion coefficient* (ADC) can be derived. When a DWI-scan is performed for at least two b-values, the ADC can be calculated with the equation 2.11, by fitting the found intensities and applied b-values to this to this equation. The acquired value for D is low when diffusion is slower, and vice versa.

2.2.4 Intravoxel incoherent motion analysis (IVIM)

The model on equation 2.11 is called mono-exponential, since it only contains a single component, describing the signal attenuation based on the diffusivity of tissue. However, it is now often suggested that the signal attenuation of DWI (and derived ADC) is also sensitive to capillary perfusion [148]. This capillary perfusion, also called the blood microcirculation, is also present in the signal, and is caused by blood flowing in a volumetric unit (voxel) [149]. This motion is

often referred to as incoherent motion, and opposes the true molecular diffusion. Since the bulk motion (flow) in capillaries is much faster than motion caused by diffusion, this microcirculation will be appearing in sufficiently low b-values, i.e. typically $< 200 s/mm^2$ [150]. Still, the signal mimics a random walk (pseudodiffusion), as the network of capillaries are randomly oriented, and causes signal attenuation in diffusion-encoding gradient pulses (e.g. PGSE sequence).

To describe both the diffusion and perfusion components, Le Bihan et al. (1986) [151] proposed a bi-exponential model:

$$I = I_0(fe^{-bD^*} * (1-f)e^{-bD}) \,, \tag{2.13}$$

that describes both compartments, with D^* being the perfusion coefficient, or pseudodiffusion coefficient and f the signal fraction between diffusion and pseudodiffusion. In the equation 2.13, information regarding perfusion can be extracted from two parameters: f and D^*. The parameter f describes the volume fraction of capillary blood flowing in a voxel. This means that f is the ratio of the volume of MR-visible water that flows in the capillary compartment, compared to the total volume of MR-visible water within the voxel. The other parameter, D^* can be viewed as a descriptor of the mimic of random walk at the scale of microcirculation. This means that D^* describes the characteristics of random motion (diffusivity) of water molecules on the macroscopic level, i.e. caused by the microcapillary flow. Therefore, when the velocity of blood flow increases, D^* increases accordingly. [149] These two parameters bear important information regarding the physiology of microcirculation of tissue.

Chapter 3

Deep learning

Deep learning (DL) is a subdomain of *machine learning* (ML) and ML itself is part of the broad concept of artificial intelligence (see figure 3.1). ML is the scientific study of algorithms that could solve tasks without specific instructions, but rather via pattern recognition. DL systems solve tasks after retrieving higher abstractions of data. To achieve this, DL models are created with functional units, or neurons, that form together artificial neural nets. This concept of DL is loosely modeled after the human brain.

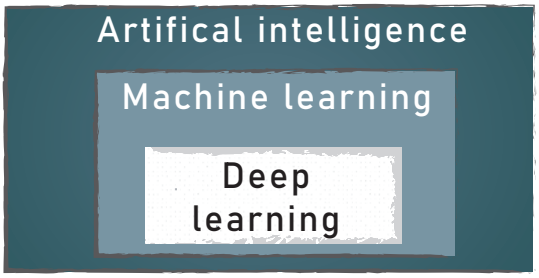


FIGURE 3.1: Schematic illustration of relations between artificial intelligence, machine learning and deep learning.

The human brain is represented as a hierarchical network of neurons. Each neuron receives an input, processes this input and pass the output to other neurons in higher layers. For the visual system, there are five basic layers of neurons: the primary visual cortex (V1), the secondary visual cortex (V2), V4 and the *inferotemporal cortex* (IT)-posterior and IT-anterior, as visualized in figure 3.2. The lower level neurons in V1 detect the basic features of the seen object, such as lines or edges. In the next depth level, i.e. V2, these found features are encoded into junctions of lines, hence patterns are detected. The V4 layer connects even more complex combinations of the features constructed in V2. In the IT-posterior, complete objects are detected, e.g. a face or a car. Finally, in the IT-anterior, more abstract or semantic meanings are connected to the found object. These principles are illustrated are summarized and exemplified in figure 3.2. [152]

In the beginning of DL and ML, neural nets were constructed of a single layer of neuron(s), e.g. *perceptrons* or *support vector machines*. Later on, more layers of neurons were added to the neural nets, and these intermediate layers were names *hidden layers*. During the past decades, facilitated by enhanced computational power, increasingly deeper networks were proposed. In the deep layers of those networks, the neurons react to abstract relations between the features found in more superficial layers. In this manner, considerably complex relations can be derived from large amounts of data.

Thus, with DL, more insights can be acquired into large amounts of data containing complex

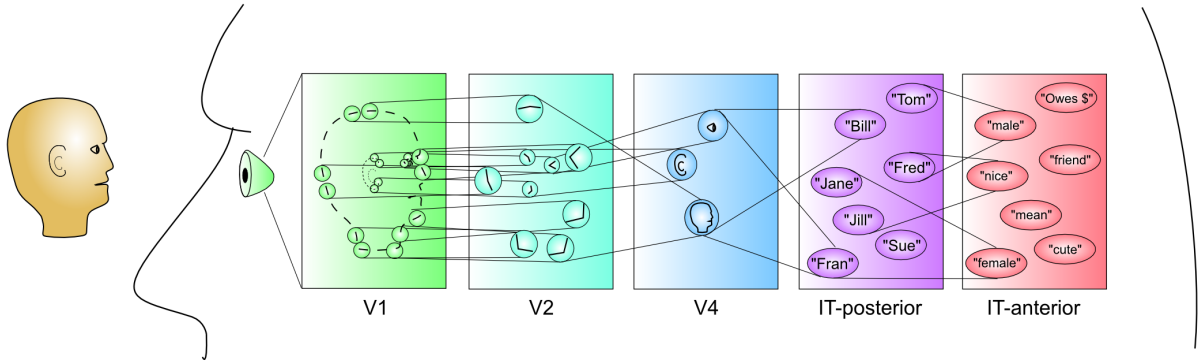


FIGURE 3.2: Schematic representation of visual system of the human brain. In basic, there exists five layers of neurons, that process a visual stimulus (V1, V2, V4, IT-posterior and IT-anterior. In analogy with DL, these layers process the input by subtracting increasingly complex features. In the last layer, more semantic meanings are connected with the found features. Figure adopted from [153].

relations. Due to the (recent) enhancement of computational power together with the availability of large amount of labeled data, DL has undergone a stunning revolution. This revolution has especially been ignited since a DL-based method won an ImageNet Large Scale Visual Recognition Competition in 2012 [154]. DL has got an important role in solving problems in a variety of terrains, including medicine, with Radiology in particular [155].

Main advantages of DL are the possibility to gain insights in complex patterns within large amounts of data, that are typically to complex to retrieve with human cognition. In medicine, this may lead to an increased efficiency, accuracy and precision, a decreased workload, increased patient face time, increased time on critical cases, costs reduction and improved monitoring [156].

3.0.1 Deep learning in Radiology

Radiology is a specialty in medicine that is upfront in adopting DL in clinical practice. It is pregnant that recently the RSNA Radiology released a journal named *Radiology: Artificial Intelligence*. It has been estimated that over 400 publications in major medical image related journals and conference venues were issued [157, 158]. Literature on DL in medical image analysis covers all parts of the field of medical image analysis, and beyond in medicine in general. DL shows excellent capabilities in classification tasks in all modalities assessing a wide range of pathologies, sometimes outperforming Radiology experts [159]. A noteworthy limitation is that these algorithms often excel in a single tasks or diagnosis, while human radiologists will diagnose way more on a medical image.

Another interesting topic of DL, also studied in the medical specialty of Radiotherapy, is the synthetic image generation. With DL-based synthetic CT, MRI-images are typically translated into CT images based on large quantities of paired real MR-CT data. Applications of these techniques are MR-only radiotherapy, but also diagnostic imaging could gain from these synthetic reconstructions. [160] Although research has been extensively conducted around the topic of DL in Radiology, a comprehensive review covering all applications is beyond the scope of this thesis.

3.0.2 Neural net's fundamentals

As stated earlier, the functional unit of a neural net is commonly called a neuron, and the equation that gives the activation a, or output, of a neuron can be defined as:

$$a = \sigma(\sum_{i=1}^{n} x_i \cdot w_i + b) \,, \tag{3.1}$$

where $x \in \mathbb{R}^n$ is the input vector, $w \in \mathbb{R}^n$ the weight vector, b the bias and σ the activation function that introduces non-linearity to the functional unit. In DL, neurons are aligned in layers to form a neural net as is shown as example in figure 3.3.

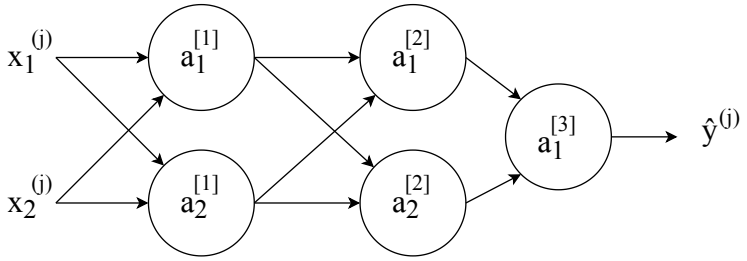


FIGURE 3.3: Schematic overview of an example network with an input layer consisting of inputs $x^{[j]}$ from training data D where $j \in D$, two hidden layers and one output layer providing the predicted class label $y^{[j]}$. The activations a^L from neurons in layer L provide the input for neurons in layer $L+1$.

Suppose labeled training data $D = (x_1, y_1), (x_2, y_2), ...(x_n, y_n)$ is available. The DL network f, composed of internal network parameters w, attempts mapping of $x \rightarrow y$ by computing $f(x, w) = \hat{y}$. The prediction error between true labels $y \in D$ and predicted labels $\hat{y}$ can be calculated consequently with $L(f(x, w), y)$. During training of f with train data D, internal network parameters w are optimized iteratively by solving:

$$\underset{w}{\operatorname{argmin}}\, L(f(x, w), y) \,. \tag{3.2}$$

This is achieved by minimizing the gradient of the loss function $\nabla L \rightarrow 0$.

The gradient of L can be found by a method called backpropagation. Partial derivatives combined with the chain rule make it possible to find the gradient for all individual weights $w_i \in w$. The chain rule is applied to find the derivative of L with respect to all individual weights, i.e. internal network parameters. Subsequently, with the found gradient for each individual internal parameter, the gradients are multiplied by the learning rate and the DL-network is updated. A single forward and backward propagation of a batch of training data is called a step, and these steps are repeated for the entire training set. A training epoch is finished when the full training set has been passed trough the network. Neural nets are typically trained with multiple epochs.

3.0.3 Convolutional neural network (CNN)

For (medical) imaging, the *convolutional neural network* (CNN) is the most popular type. A CNN is a subclass of DL, and does typically exist of convolutional kernels instead of neurons. A convolution is a mathematical operation between two functions (f and g), creating a third function ($h = f * g$) that describes one function modified by the other. In the context of DL, convolutions are powerful in subtracting visual features from imaging data.

A discrete convolution filter is a matrix (kernel) that filter features from an input image, by a point-wise multiplication across the input image, creating an output feature map. So, suppose f is a 2D medical image and g is a convolution kernel with dimensions, the feature map h becomes at row m and column n:

$$h[m,n] = f[m,n] * g[m,n] = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} f[i,j] \cdot g[m-i, n-j] \,. \tag{3.3}$$

With these convolutions, increasingly complex features can be derived from an image. Often, an activation layer is applied, and with pooling functions the dimensions are reduces and features are extracted from the convolution outputs. Optionally, conventional artificial neurons are applied after the convolutions layers, to connect features in a spatial independent manner and return the final output of the network, e.g. a classification. Besides classifications, image segmentation and synthesis are tasks often solved with CNNs [157].

Generative adverserial network (GAN)

A specific subtype of DL and CNNs are the *generative adverserial networks* (GANs). GANs have gained a considerable amount of attention, mainly due to their capabilities in generation of data. With GANs, two neural networks are trained simultaneously. One of the two networks focuses on generation of data, while the other network has the task to discriminate the generated data from the real data in the training set. GANs have achieved high performance in several tasks in data generation, including text-to-image synthesis, super-resolution and image-to-image translation [158]. A more detailed description and fundamentals on the network architecture of a GAN is given in section 6.1.3.

Chapter 4

Problem statement

4.1 Problem statement

MR-HIFU is an attractive treatment option for symptomatic UF, due to its safety, non-invasive character and effectiveness. Together with uterine fibroid embolization, it offers a non-surgical treatment option for UF that is uterus-sparing.

Postprocedurally, the treatment effect of MR-HIFU is standardly assessed with a CE-T1w MR-scan, where a high NPV ratio indicates a successful treatment. However, a drawback of the MR-HIFU procedure is the inability to visualize the treatment effect intraprocedurally. Risks of Gd entrapment and distorted thermometry inhibit the use of contrast-agents between MR-HIFU sonications, making further ablation after administration of a contrast agent unsafe. This pitfall commonly entails long procedure times and limited NPVs, caused by either delayed or premature ending of sonications, as established on a post-procedural CE-T1w scan when administration of a Gd-based contrast agent is safe.

To undo dependency on Gd-based contrast agents and thereby enable intraprocedural NPV assessment, efforts have been made to evaluate the MR-HIFU treatment effect based on multiparametric MRI. It has been suggested that extraction of parameters from DWI scans could be helpful in discriminating the NPV from the perfused volume. Low b-values particularly relate to perfusion characteristics of UF tissue, and could therefore play a pivotal role in in Gd-free visualization of the MR-HIFU treatment effect. Until now, previous work about DWI has only been analyzed with a mono-exponential ADC model in the context of periprocedural NPV visualization.

4.2 Thesis

The non-perfused volume of uterine fibroids after MR-HIFU treatment can be assessed adequately without a gadolinium-based contrast agent by analysis of diffusion weighted imaging with *intravoxel incoherent motion* modeling or deep learning.

4.3 Approach outline

The aim of this work was to further explore DWI-based methods to visualize the progression of MR-HIFU ablations peri- or intraprocedurally, thus avoiding the need for a contrast agent to assess the NPV. The first three introductory chapters of this work described the general background of the topic, by summarizing the existing literature. In this thesis, two approaches are investigated to visualize the NPV by using DWI: IVIM and a DL-based method. At first, with the IVIM strategy it was attempted to describe the underlying physiological principles of UF perfusion with a comprehensible model. Thereafter, the DL-based method was applied to potentially

achieve better results for the stated problem, but at the cost of understanding the relation of the DWI data with the applied model. The methods and results of both techniques are described sequentially in chapter 5 and chapter 6. Finally, a general discussion is presented reviewing on both methods.

Chapter 5

Approach 1: IVIM

This chapter covers the first approach for using DWI to avoid the need for a gadolinium-enhanced contrast agent to visualize the NPV during or after a MR-HIFU treatment of UF. This first approach will be *intravoxel incoherent motion analysis* (IVIM) , and in the first first section an outline of the method is given, followed by the obtained results in the second part.

5.1 Methods

5.1.1 Patients

This study made retrospectively use of the single-center dataset described in the prospective multi-parametric MR-study MaSSII-study from Verpalen et al. (2020) [161]. These patients were included from January 2018 to January 2019 at the Isala Hospital in Zwolle, The Netherlands. At first, the patients were screened for eligibility by a gynaecologist. The in- and exclusioncriteria are listed in table 5.1.

Inclusion criteria	• Age of 18 years • US-confirmed diagnosis of UF • pre- or peri-menopausal state
Exclusion criteria	• Current pregnancy • MRI contra-indications • Calcified UF • Fat layer > 4 cm • Concomitant adenomyosis • Supsicion of malignancy • No contrast-enhancement of the dominant UF on the T1w image • Maximum diameter of the dominant UF < 1cm • > 10 UFs • Fibroid signal intensity on the T2w image and contrast-enhanced T1w image higher compared to the signal intensity of the myometrium • Inaccessible dominant UF

TABLE 5.1: Eligibility criteria for the MaSSII-patients with UF for MR-HIFU treatment.

The study from Verpalen et al. (2020) described a total of 56 included women with a DWI sequence. All these patients received a screening MR-scan, and patients eligible for MR-HIFU underwent MR-HIFU treatment of UF(s) and a six months follow-up MR-scan. In this work, patients that did not receive an MR-HIFU treatment were excluded. The characteristics of the remaining patients with suitable MR-datasets are listed in table 5.2.

Characteristic		Mean or count
Patients included		56
Age (years)		42.8 ($\pm$ 7.0)
BMI (kg/m^2)		24.8 ($\pm$ 3.6)
Abdominal fat layer (cm)		2.1 ($\pm$ 1.2)
Uterine position	AVF	45 (80.4)
	RVF	6 (10.7)
	Upwards	5 (8.9)
Number of fibroids treated	1	26 (46.4)
	2	11 (19.6)
	3	7 (12.5)
	4	2 (3.6)
	5	4 (7.1)
	>5	6 (10.7)
Fibroids' location	Submucosal	22 (26.8)
	Intramural	23 (28.0)
	Subserosal	22 (26.8)
	Hybrid	15 (18.3)
Scaled Signal Intensity		12.0 ($\pm$ 14.5)
Funaki class	1	8 (9.8)
	2	64 (78)
	3	10 (12.2)
Maximum diameter (cm)		5.53 ($\pm$ 3.4)
Fibroid volume pre-HIFU (cm)		132.9 ($\pm$ 220.6)
Non-Perfused Volume (%)		64.8 ($\pm$ 34.3)

TABLE 5.2: Characteristics of the retrospective dataset from the included MaSSII-study (adopted from Verpalen et al. (2020) [161]) The quantities are expressed in mean or counts (% or $\pm$ SD).

5.1.2 Scan protocol

As stated above, the data was collected in a single-center at the Isala Hospital in Zwolle. The MR-HIFU treatments were carried out on a clinical HIFU system (Sonalleve, V1, Profound Medical Inc, Mississauga, Canada). This HIFU system contains a tabletop transducers, that can be integrated with the 1.5-T MR scanner (Achieva, Philips Healthcare, Best, The Netherlands). MR-scans were acquired during screening, treatment and follow-up. For all these moment, the same DWI-protocol was used and included a scan with TE=65ms and 7 b-values (i.e. DWI weightings), with 0, 50, 100, 200, 400, 600 and 800 s/mm^2, with an example visualized in figure 5.1. The DWI-scans were acquired prior to contrast injection. At these moments in time, a T2w scan was also acquired for identification of anatomical structures, and the protocol was ended after acquisition of a CE-T1w scan. This CE-T1w was a 3D spoiled gradient echo scan, and used for viability assessment during screening, and NPV visualization post-HIFU. The administered contrast agent was gadolinium-based (DOTAREM, 0.2 mL/kg, Gadoterate Meglumine, 0.1 mmol/kg, Guerbet; Aulnay-sous-Bois, France). An example 3D axial DWI-slice (x,y,b) is depicted in figure 5.1

5.1.3 Preprocessing

Let $\vec{D}$ be a 4D DWI volume with m rows, n columns, p layers and q DWI-weightings (b-values):

$$\vec{D} = (a_{ijkb}) \in \mathbb{R}^{m \times n \times p \times q}, \tag{5.1}$$

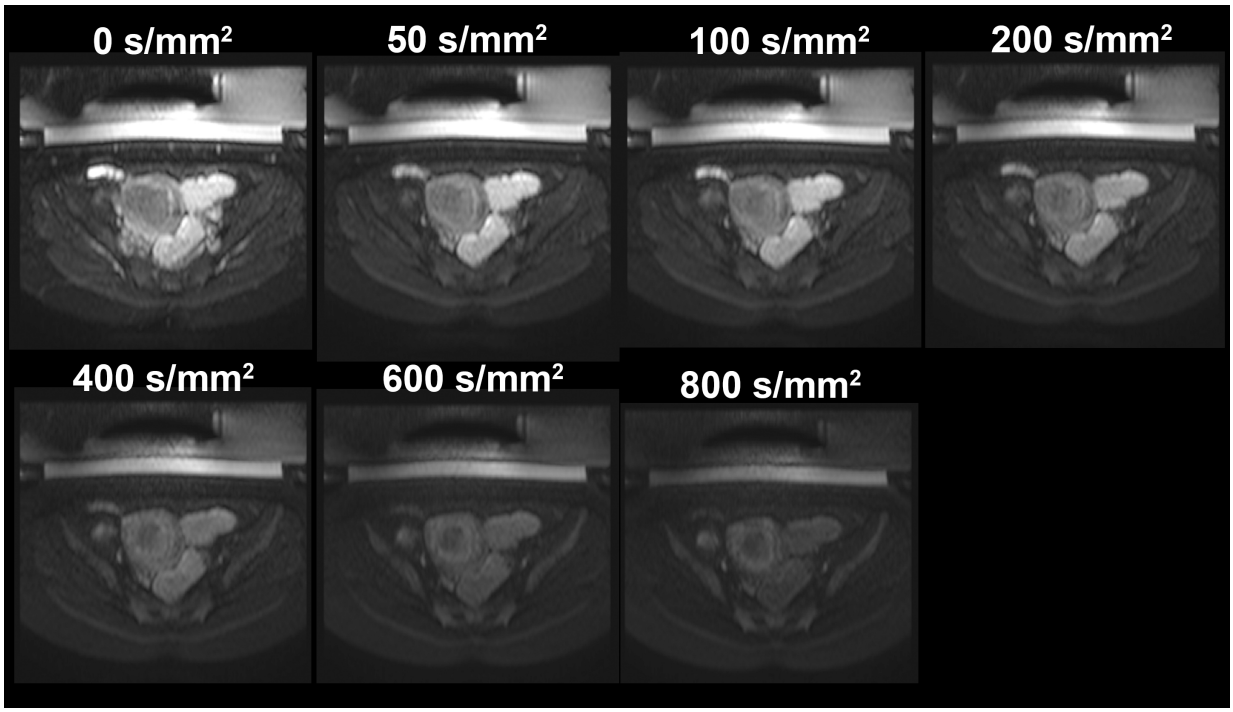


FIGURE 5.1: Example slice of an axial DWI-scan of a female pelvic region with the uterus bearing an UF that received MR-HIFU treatment. The DWI datasets processed in this work included the b-values depicted here.

where i,j,k denote the entry a (i.e. voxel) at the i-th row, j-th column and k-th layer respectively in $\vec{D}$, and b the complementary b-value of the volume ($b \in B$).

Registration of b-values

Registration of the b-values is performed to minimize the differences in the spatial domain of all volumes along the axis of the DWI-weightings in $\vec{D}$. This registration involves the problem of finding an displacement vector u_{ijkg} that spatially aligns a_{ijkg} to a_{ijkb_0}, with g being a b-value in B, excluding $b = 0s/mm^2$, and a an entry in $\vec{D}$:

$$a_{ijkg} + u_{ijkg} \xrightarrow[\text{aligns}]{} a_{ijkb_0} \, . \tag{5.2}$$

For registration the *Elastix*-toolbox [162] is used for registration of the non-zero b-values to the b_0 DWI volume.

Masking and normalization

After registration and before performing the IVIM fit, all volumes were masked by the 99-percentile of $\vec{D}$ based on the b_0 volume:

$$\vec{D}_{masked} = \begin{cases} 0, & \text{if } a_{ijkb_0} < T \\ a_{ijkb}, & \text{otherwise} \, . \end{cases} \tag{5.3}$$

where T the 1-th percentile of $\vec{D}_{i,j,k,b_0}$. Secondly, the masked volumes are normalized to an average pixel intensity of 100 at the b_0 slice:

$$\mu = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}\frac{a_{ijkb_0}}{m+n+p}\,, \tag{5.4}$$

$$\vec{D}_{normalized} = \frac{\vec{D}_{masked}}{\mu} * 100\,, \tag{5.5}$$

With $\vec{D}_{normalized}$, the final form of the preprocessed matrix $\vec{D}$ is achieved, that is used for IVIM analysis.

5.1.4 IVIM fit

With the IVIM model, the relation between b-values and signal intensity I is described with the following bi-exponential equation (same as 2.13:

$$I(b) = I_0(fe^{-bD^*} + (1-f)e^{-bD})\,, \tag{5.6}$$

with the parameters:

- I_0: initial signal intensity (at b_0, without diffusion weighing)
- f: signal fraction of pseudodiffusion
- D^*: pseudodiffusion coefficient
- D: diffusion coefficient

As explained in the introduction of this work: equation 5.6 describes two compartments, with two hypothesized underlying physiological principles: pseudodiffusion, or perfusion, and diffusion. The parameter f reflects the volume fraction of both compartments present in the signal intensity I.

5.1.5 Fit method

The parameters from equation 5.6 are retrieved from the individual voxels by performing a voxel-wise fit to the 4D DWI volumes. The voxel-wise fits are executed with a *least-squares* optimization algorithm. A least-squares method is often used in regression problems, and aims at minimizing the sum of the squared residuals, i.e. the difference between the observed values and values provided by the model used for fitting.

Applied to the equation 5.6, if β is a vector containing the IVIM parameters, a residual at b-value b of a given voxel $a \in \vec{D}$ is defined as:

$$r(b,\beta) = a_b - I(b,\beta)\,, \tag{5.7}$$

where b is an element of B (see section 5.1.2). The optimal parameters β for a voxel are found by minimizing the sum of the squared residuals S, for $b \in B$:

$$S(\beta) = \sum_{b=0}^{800} r(b,\beta)^2\,. \tag{5.8}$$

A minimum of S is achieved by setting the gradient to 0. Here, the *Trust-region-reflective* (TRF) algorithm is used to achieve minimalization of S, and is explained the appendix, section A.1.

5.1.6 Initialization fit

For each preprocessed volume $\vec{D}$, a voxel-wise fit with equation 5.6 as objective function and TRF as fit algorithm is performed. To perform voxel-wise fitting by using the TRF algorithm, defining values for β_0 that approaches the minimizer is important. Therefore, a initialization fit is performed based on b-values of the average of all voxels $a_{mean,h}$ to find values for the β_0 vector as initialization of the voxel-wise fit. Also for this initialization fit, choosing a adequate starting point is important. The following values are chosen as β_0 for the initialization fit and based on existing literature [147, 149, 151, 163]:

$$\beta_{0_{init}} = \begin{bmatrix} I_{0_{init}} \\ F_{init} \\ D_{init} \\ D^*_{init} \end{bmatrix} = \begin{bmatrix} 100 \\ 0.2 \\ 1.0 * 10^{-3} \\ 1.0 * 10^{-2} \end{bmatrix} . \tag{5.9}$$

Please note the difference between 4D DWI volume $\vec{D}$, and the parameter describing diffusivity D (without vector symbol). The TRF algorithm takes the mean values of all voxels in $\vec{D}$ for each b-value, the IVIM equation (equation 5.6) and β_0 as input, and delivers the final β_f as output, when a certain threshold is reached:

$$TRF(a_{mean,b}, I(b), \beta_{0_{init}}) \xrightarrow{fit} \beta_0 = \begin{bmatrix} I_{0_0} \\ F_0 \\ D_0 \\ D^*_0 \end{bmatrix} . \tag{5.10}$$

After the algorithm is finished, the β_0 are defined and used for the voxelwise fit. During the fitting, the search space is constrained by the following bounds:

$$\begin{aligned} 0 < I_0 < \infty \\ 0 < F < 1.0 \\ 0 < D < 3.5 * 10^{-3} \\ 0 < D^* < 5.0 * 10^{-1} . \end{aligned} \tag{5.11}$$

The TRF fits are performed using the least squares implementation of *SciPy* [164].

5.1.7 Voxel-wise fit

At each index i, j, k in 4D DWI volume $\vec{D}$ the parameters β are approached by using the TRF algorithm, starting at β_0 found in the initialization fit. In order to decrease the influence of noise on the parameter estimation, the number of parameters is reduced by fixing the D^* parameter:

$$I_2(b) = I_0(fe^{-bD^*_0} + (1-f)e^{-bD}) , \tag{5.12}$$

This means that the degrees of freedom during the search is limited, and the fitting process for voxels at indices i, j, k becomes:

$$TRF(a_{ijkb}, I_2(b), \beta_0) \xrightarrow{fit} \beta_{ijk} = \begin{bmatrix} I_{0_{ijk}} \\ F_{ijk} \\ D_{ijk} \\ D^*_0 \end{bmatrix} . \tag{5.13}$$

When for each voxel at entry i, j, k is the fitting parameters are estimated, the values of the parameters at each entry constitute a map, e.g. an F-map, D-map or I_0-map. For the voxel-wise

fits, the same constraints are used from equation 5.11 and implementation of *SciPy* [164]. For both fits (initialization and voxel-wise), the search was terminated after a change of the cost, fit parameters, or gradient smaller than threshold value $1.0 * 10^{-8}$; or when the number of function evaluations exceeded 100 times the number of fit parameters.

5.1.8 Evaluation

The ability of this method to distinguish the perfused from the non-perfused volume is assessed quantitatively based on the resulting parameter-maps (F,D,I_0) from the entire dataset. This is investigated by quantitative examination of contrast differences between the anatomical areas of interest. The quantitative method to detect whether statistically significant contrast differences between pixel intensities in the different areas of interest, is elaboarted in sections 5.1.8 and 5.1.8. To define which voxels are in the specific areas of interest, ROIs are drawn as listed in table 5.3 and illustrated in figure 5.2. All ROIs were created by a radiology resident and a medical student trained in MRI reading. Since the FV and NPV regions are drawn, the PV is defined mathematically as:

$$PV_{ROI} = FV_{ROI} - NPV_{ROI} = \{x | x \in FV_{ROI} \wedge x \notin NPV_{ROI}\}\,, \tag{5.14}$$

where x is a element inside an ROI. Thus, the PV is the difference between the FV and NPV ROIs. In this chapter, the contrast for the parameters D and f are evaluated based on these ROIs.

moment in time	Region	ROI based on
screening	uterine fibroid	screening T2w
post-HIFU	uterine fibroid	pre-HIFU T2w
post-HIFU	NPV	post-HIFU CE-T1w
post-HIFU	PV	post-HIFU uterine fibroid and NPV ROI

TABLE 5.3: ROIs used for evaluation in this chapter, and chapter 6.

NPV contrast

An important aspect of the parameter maps is the contrast between PV and NPV. This contrast determines the ability to separate the NPV from the PV, and is therefore of great clinical importance to visualize the NPV. Contrast is defined here as the difference in average parameter value between the NPV, the PV_{pre} and the $FV_{screening}$. Therefore, the average pixel intensities per parameter map are calculated and reported. In addition, the intensities between the different ROIs are tested for statistical significance for each parameter map, as explained below, in section 5.1.8.

In addition, the intensities of the $f-$ and D-maps will be analyzed on contrast differences between the PV and NPV on individual patient level. The fraction of patients where significantly different contrast is present will be reported in percentage. This analysis on individual patient level is only performed for NPV and PV ROIs, since those are deterministic for the ability to separate the ablated volume from viable tissue.

Statistical analysis

The parameters were expressed as mean $\pm$ standard deviations. The means were tested for normal distribution using the D'Agostino's K-squared test with 5%: $\alpha = 0.05$ [165, 166]. To detect whether mean values of the NPV and the PVs from the post-HIFU and screening reconstructions differ significantly, the Friedman test was applied between the means of D and f of the

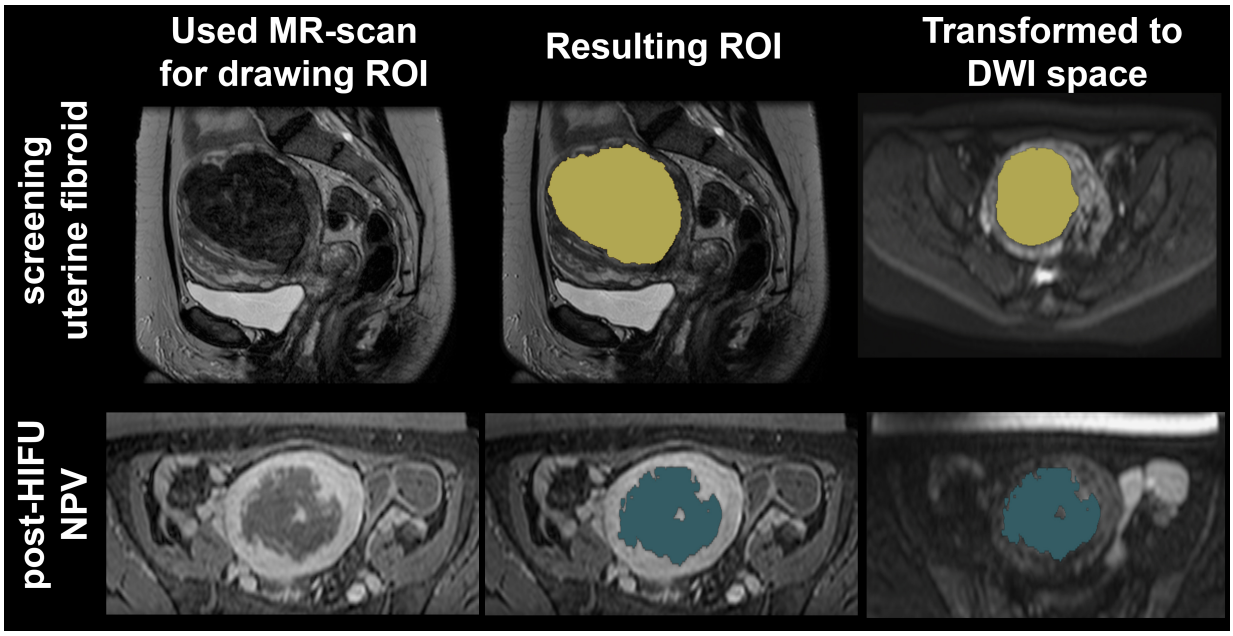


FIGURE 5.2: Example of drawn ROI. T2w MR-scans were used for delineating the FV, while the ROIs for the NPV were drawn on the CE-T1w MR-scans. It should be noted that for both the screening and direct pre-HIFU T2w MR-scans ROIs of the FV were composed. The PV can be derived from the difference in ROI between the pre-HIFU FV and the NPV. The UF visible here has Funaki score 2, FIGO class 4 and an NPV-ratio of 59.6% after MR-HIFU treatment.

three groups with a significance level of 5%: $\alpha = 0.05$. For *post-hoc* analysis when the Friedman test returned a p-value < 0.05, the Nemenyi test was executed to determine the underlying differences in distribution of the regions [167].

For testing contrast differences on patient level from the f- and D-map within the PV and NPV ratios, a Mann-Whitney U-test was used with $\alpha = 5\%$.

5.2 Results

First, some graphical examples are shown from individual patients, in second, overall results are presented.

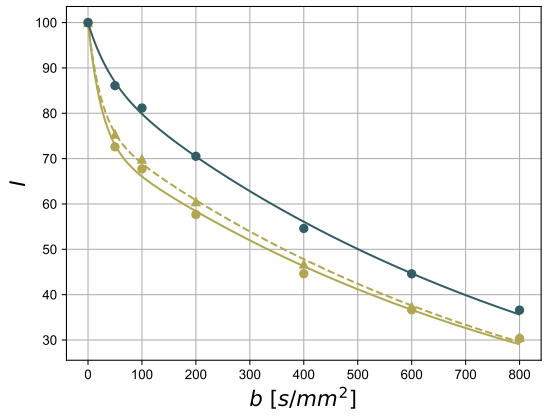
5.2.1 Individual examples

In figure 5.3 the results of the IVIM fit are plotted from the patient visible in figure 5.2, with the depicted ROIs. The mean measured DWI datapoints are graphed, with the lines constructed with the parameters acquired after the fitting with the IVIM equation (5.6). The same data is illustrated on a linear scale and a logarithmic scale, in figure 5.3 (a) and (b) respectively.

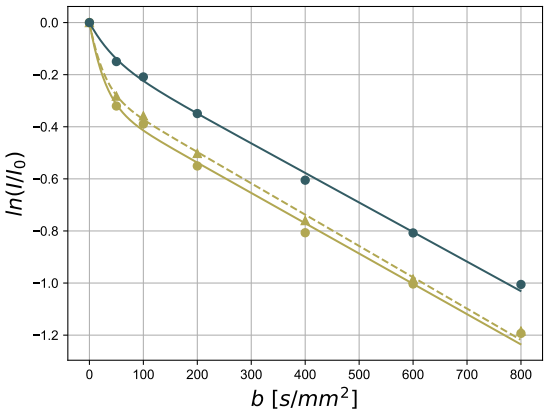
Next, visual parameter maps are shown in figure 5.4. In this illustration, complementary T2w and CE-T1w slices are given, together with the parameters derived from IVIM fitting, i.e. the I_0-, D- and f-maps.

5.2.2 Overall results

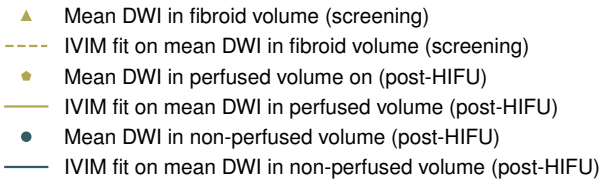
The mean values of D and f for the combined groups were $1.08 * 10^{-3}(\pm 2.22 * 10^{-4})$ mm^2/s and $2.12 * 10^{-1}(\pm 7.57 * 10^{-2})$ (no unit) respectively. The average D-parameter values were $1.25 *$



(a)



(b)



(c)

ROI	$D[mm^2/s]$	$D^*[mm^2/s]$	f
FV (screening)	$1.20 * 10^{-3}$	$4.19 * 10^{-2}$	$2.26 * 10^{-1}$
PV (post-HIFU)	$1.16 * 10^{-3}$	$4.02 * 10^{-2}$	$2.63 * 10^{-1}$
NPV (post-HIFU)	$1.14 * 10^{-3}$	$2.41 * 10^{-2}$	$1.16 * 10^{-1}$

(d)

FIGURE 5.3: Example IVIM plots from patient visualized in figure 5.2, with DWI data inside the ROIs as illustrated after a successful MR-HIFU treatment. (a) IVIM plots on linear intensity (I) scale. (b) IVIM plots on logarithmic scale ($ln(I/I_0)$). (c) Legend for the graphed data in both scales. (d) Parameters retrieved from IVIM fitting, and used for calculation of the IVIM curves in this image.

$10^{-3}(\pm 4.30 * 10^{-4}\ mm^2/s)$, $1.01 * 10^{-3}(\pm 4.47 * 10^{-4}\ mm^2/s)$, and $1.13 * 10^{-3}(\pm 4.46 * 10^{-4}\ mm^2/s)$ for the FV, PV and NPV respectively. For the f-parameter, the mean values of the FV, PV and NPV were $0.26(\pm 0.19)$, $0.29(\pm 0.21)$, and $0.12(\pm 0.13)$ respectively.
In figure 5.5 the distribution of the D and f parameter is graphed in a boxplot for all the three separate regions (i.e. screening FV, post-HIFU PV and post-HIFU NPV) in all patients. The distribution of the mean parameters in the aforementioned regions is plotted in figure A.4. This graph contains a scatterplot of the D versus the f parameter. In table 5.4 are the results listed from the D'Agostino's K-squared test. Except for the parameters in the post-HIFU PV region and the D in screening FV, the parameters did not match a normal distribution ($p > 0.05$).

The outcome of the Friedman test statistic for the D parameter is $1.03 * 10^1$ with a p-value of $5.81 * 10^{-3}$. For the f-parameter, the test results in a statistic of $3.23 * 10^1$ and a p-value of $9.7 * 10^{-8}$. As the null-hypothesis is rejected for both the D and f parameter, Nemenyi *post-hoc* analysis is conducted. The p-values of the *post-hoc* test is tabled in table 5.5, and shows that for the D parameter, a statistically significant difference exists between the region of the FV during screening, and the PV after MR-HIFU treatment. No other null-hypotheses can be rejected for the D coefficient. For the f parameter, the difference between the mean values of the NPV and the PV post-HIFU is statistically significant, together with the difference between the NPV and the FV during screening.

For contrast differences between the NPV and PV of the individual patients, statistically significant difference in pixel intensities was found for both the f- and D-parameter maps in 89% of patients. For f, in all 89% of cases, mean intensity in the NPV region was lower than

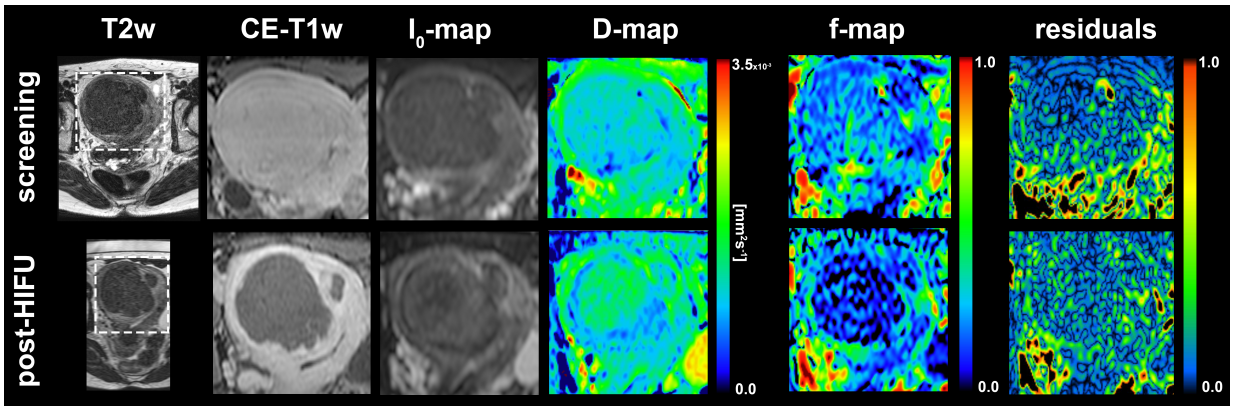
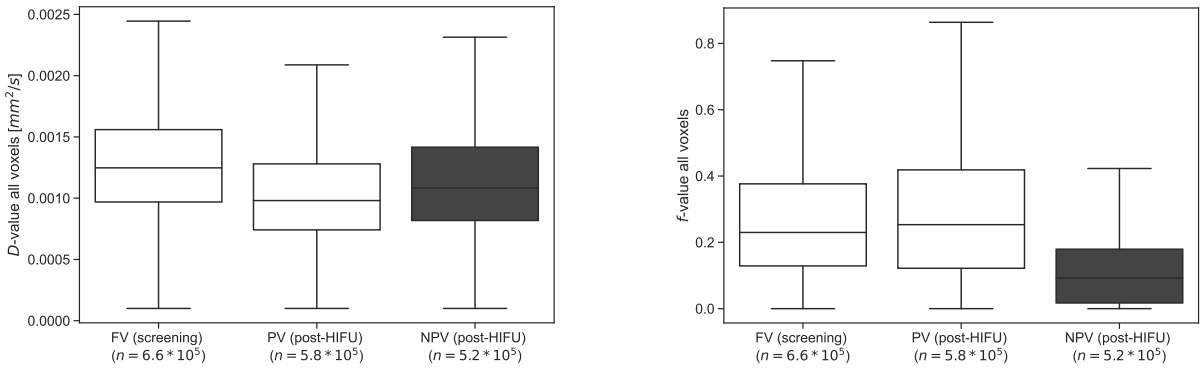


FIGURE 5.4: An example of an IVIM reconstruction before and after MR-HIFU treatment where near-complete ablation was achieved for two UFs (263 and 14 mm^3), with NPV-ratios of 84% and 40% respectively. For both moments in time, a T2w and CE-T1w scan are depicted, together with maps derived from the IVIM analysis (i.e. $I_0-, D-$ and f-map). Also, the residual map is shown, where reconstruction errors are visualized after recalculation of DWI intensities with the extracted IVIM parameters, via equation 5.6 and b-value 50 s/mm^2.

within the PV. In case of D, from the 89% of patients with statistical differences in intensity between NPV and PV, 65% was on average higher, and 35% lower. This indicates that contrast differences between the NPV and PV are more consistent for the f-parameter.

Region	D		f	
	k	p-value	k	p-value
FV (screening)	0.699	0.705	8.65	0.0132
PV (post-HIFU)	2.27	0.321	0.707	0.702
NPV (post-HIFU)	7.68	0.0215	14.8	$6.10 * 10^{-4}$

TABLE 5.4: D'Agostino's K-squared test for the distribution of the f- and D-parameter extracted with the IVIM analysis in the different ROIs. The goal of this analysis was to test the following null-hypothesis: *parameter x is sampled from a normal distribution*, where x is either f or D.



(a) (b)

FIGURE 5.5: Overall results of all the cases from the retrospective DWI datasets. The parameters were extracted with the IVIM equation (5.6), and the mean values are visualized in these boxplots. The voxels are separated for the different three ROIs, in order to investigate whether differences in distributions exists between these regions. (a) Distribution of D visualized in a boxplot of all voxels in the DWI datasets in the three different regions before and after MR-HIFU treatment of UF. (b) Boxplot visualization of the distribution of the f parameter of all voxels in the three regions. Note that the large number of datapoints, i.e. voxels, result in a large number of outliers in these boxplot, caused by noise within the DWI data.

		FV (screening)	PV (post-HIFU)	NPV (post-HIFU)
FV (screening)	D	-1.00	0.011	0.18
	f	-1.00	0.57	< 0.001
PV (post-HIFU)	D	0.011	-1.00	0.50
	f	0.57	-1.00	< 0.001
NPV (post-HIFU)	D	0.18	0.50	-1.00
	f	< 0.001	< 0.001	-1.00

TABLE 5.5: p-values for the Nemenyi *post-hoc* analysis after rejection of the null hypothesis with the Friedman test ($\alpha = 0.05$). These outcomes are visualized in figure 5.6.

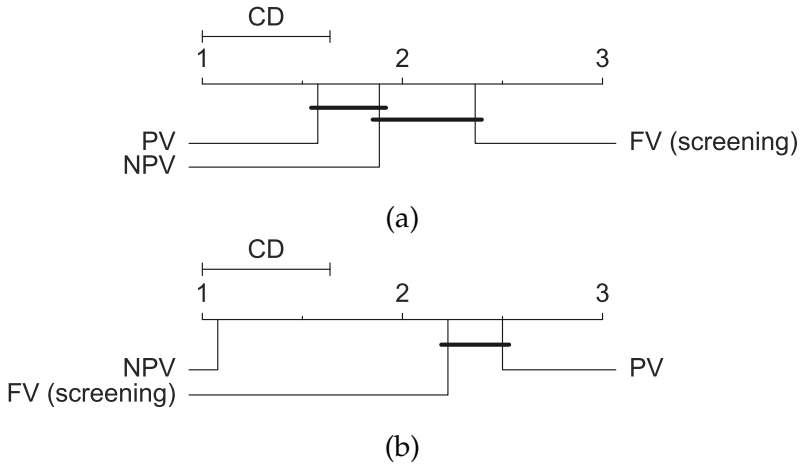


FIGURE 5.6: Visual presentation of the *post-hoc* analysis after the Friedman test, with the *critical difference* (CD) calculated with the Nemenyi test. Distances between ranks greater than CD are considered statistically significant different. The horizontal axis represents the mean rank of the intensities within the different regions. Mean ranks that are not significantly different are connected with the black bar. (a) The diagram for D coefficient. It can be seen that for this parameter, the PV (post-HIFU) and FV (screening) are not connected, thus, statistically significant different. (b) the diagram for f coefficient. Since the NPV (post-HIFU) is not connected with FV (screening) and PV (post-HIFU); the pixel intensities within the NPV (post-HIFU) ROI differ significantly from the intensities in the FV (screening) and PV (post-HIFU) ROI, based on the f-map.

Chapter 6

Approach 2: DL-based

The result that certain parameters extracted from the DWI-data bear a relationship with the size and location of the NPV. Since the relationship between the DWI-data and the NPV could be more complex than is described with the IVIM equation (equation 5.6) and might be spatially dependent. Hence, a DL-based approach could potentially be fruitful. This means that the regression is not performed with a set equation, but with a much larger amount of fitting parameters by composing a neural net. This approach is explored in this chapter by using a cGAN.

6.1 Methods

6.1.1 Data and subsets

Also this approach is based on DWI-data, and used the same dataset as described in section 5.1.1 and 5.1.2. For DL, division of the dataset in unique subsets is important for developing, validating and testing of the performance of the algorithm. Here, the total dataset D_{tot} containing all patients was divided on patient level in a training (D_{train}) validation (D_{val}) and testing (D_{test}) subset on a 0.4-0.2-0.4 ratio respectively. There was no overlap between the sets, mathematically this means:

$$\begin{cases} D_{train} \cap D_{val} \\ D_{train} \cap D_{test} \\ D_{val} \cap D_{test} \end{cases} = \emptyset \,, \tag{6.1}$$

where $\emptyset$ is an empty set. After division, the following DWI-volumes were excluded:

- DWI volumes with a non-standard field-of-view (i.e. field-of-view only covering UF, instead of whole axial body,
- DWI volumes with a low spatial resolution.

6.1.2 Preprocessing

At first, the 3D ground-truth CE-T1w scans were spatially aligned to the DWI-datasets, based on the b_0 volume:

$$e_{ijk} + u_{ijk} \xrightarrow[\text{aligns}]{} a_{ijkb_0} \,, \tag{6.2}$$

where e is an entry from a CE-T1w-scan, u is the resulting transformation vector, and a an entry from the matching DWI-scan. Registration was performed using the *Elastix*-toolbox [162], with a intensity-based B-spline registration algorithm with an adaptive stochastic gradient descent optimizer. Secondly, both the DWI and CE-T1w were normalized to the 99-th percentile, as described in equations 5.4 and 5.5. No clipping was applied, contrary to the method described at

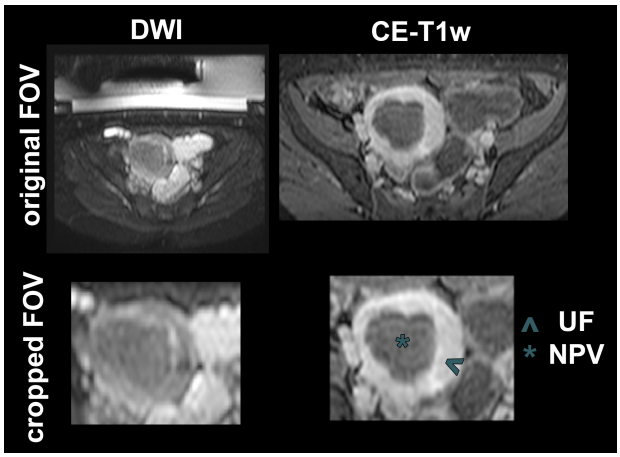


FIGURE 6.1: A DWI and CE-T1w slice with an original Field-of-view (FOV) of the female pelvic anatomy with an UF, for the DWI and true CE-T1w scan (upper row). The lower row depicts the subsequently cropped FOV, as fed to the neural network.

the IVIM chapter (5.3).

Thereafter, due to the UF being the region of interest in this problem, surrounding tissues were excluded during training. This was realized by cutting out the outer parts by drawing a cube around the uterine area, while preserving the uterus including the UF. An example of cropping is illustrated in figure 6.1.

Lastly, the 3D CE-T1w and 4D DWI volumes were reduced to 2D and 3D respectively by axial slicing:

$$\vec{D}_k = (a_{ijkb}) \in \mathbb{R}^{m \times n \times 1 \times q} \tag{6.3}$$

$$\vec{C}_k = (e_{ijk}) \in \mathbb{R}^{m \times n \times 1} \tag{6.4}$$

where $\vec{D}$ is a DWI-scan and $\vec{C}$ a CE-T1w-scan and k denotes the index of the axial slice. An example DWI-slice is visualized in figure 5.1. The cropped slices were made a square by applying zero padding to the smallest dimension, and afterwards rescaled to a resolution of 128×128 pixels. The DL-network is trained on these patches, and the volumes are recreated with a patchwise reconstruction with a stride of 4 pixels.

6.1.3 Network architecture

The main network architecture is based on an image-to-image translator with a conditional-GAN. Here, the implementation of *pix2pix* [168] forms the basis of the networks deployed in this project, and an overview is illustrated in figure 6.2.

Generator

The generator is composed with an encoder-decoder shape with skip connections, often referred to as *U-Net* [169]. The feature maps of the filters during the encoding are propagated via the skip connections to the decoding part, with concatenation. This makes it possible that during the expansion of the features from the latent space (the features in the bottleneck), the high resolution features retrieved during contraction can be utilized as well. The encoder of this architecture consists of the following number of filters for each layer: $64 - 128 - 256 - 512 - 512 - 512 - 512$. The decoder contains $512 - 512 - 512 - 256 - 128 - 64$ filters. Note that these numbers represent

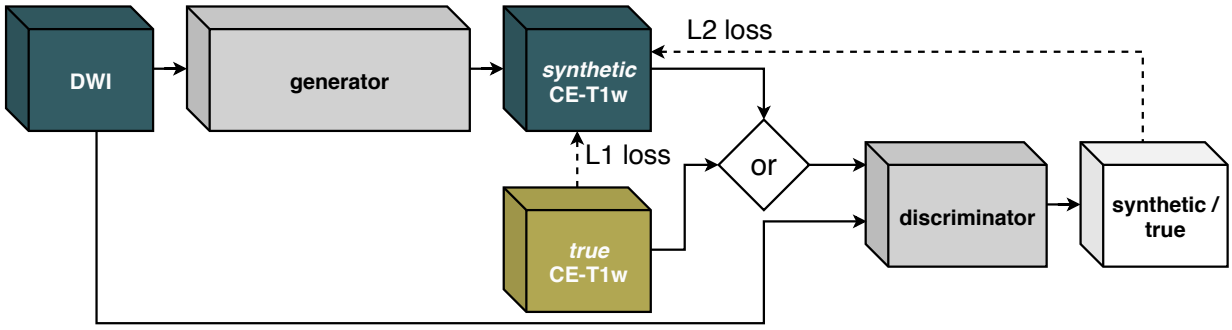


FIGURE 6.2: Schematic overview of the used cGAN during this project. With a DWI dataset and the generator a synthetic CE-T1w image is created. The discriminator's goal is to distinguish the synthetic from the true CE-T1w scan. During training, the generator and discriminator try to outperform each other. In this process, the generator's internal parameters are updated with an $L2$ loss from the inverted discriminator, and an $L1$ loss from the true CE-T1w. The generator and discriminator components are explained in detail in figure 6.4 and 6.5.

the amount of output feature maps per layer, instead of input feature maps.

The output of the convolutional layers in the generator are followed by a *LeakyReLu* delinearization function with a slope of 0.2 (see figure 6.3).

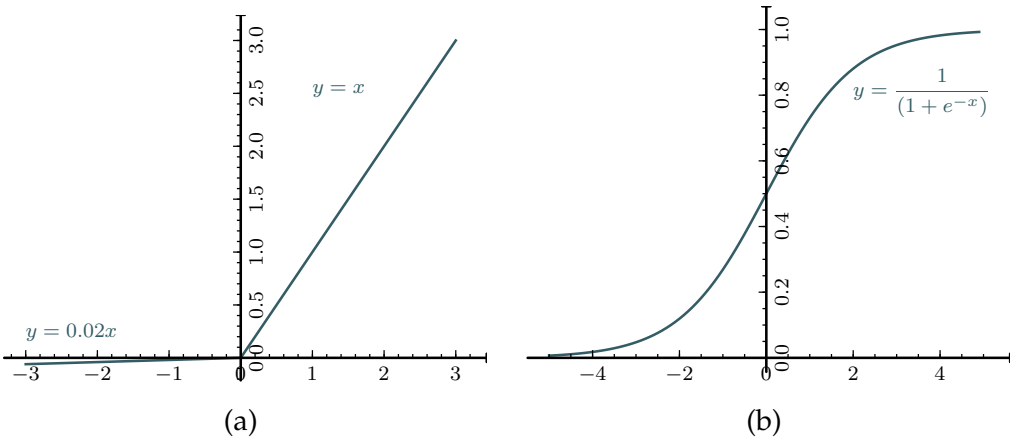


FIGURE 6.3: Activation functions used for delinearization of the output of the generator layers. (a) LeakyReLu with a slope of 0.02 is used here as activation function of the convolutional layers. (b) Sigmoid delinearization as activation function for the output of the last convolutional layer.

The activation of the layers are completed with a *BatchNormalization* (BN) function, that normalizes the activation at each batch. This means that with BN the activation of each batch is maintained at a mean of 0 and a standard deviation of 1, by scaling with a learnable parameters ζ and shifting factor γ:

$$y_i \leftarrow \zeta x_i + \gamma \equiv BN_{\gamma,\zeta}(x_i) \tag{6.5}$$

at the i-th layer, where x is the intermediate activation before BN, and y the activation after BN is applied. This means that each layer has two extra learnable parameters that help converging the network. With BN, the moving average during training is taken into account when updating the BN parameters, and is called momentum (weighing factor). Here, momentum is set to 0.8.

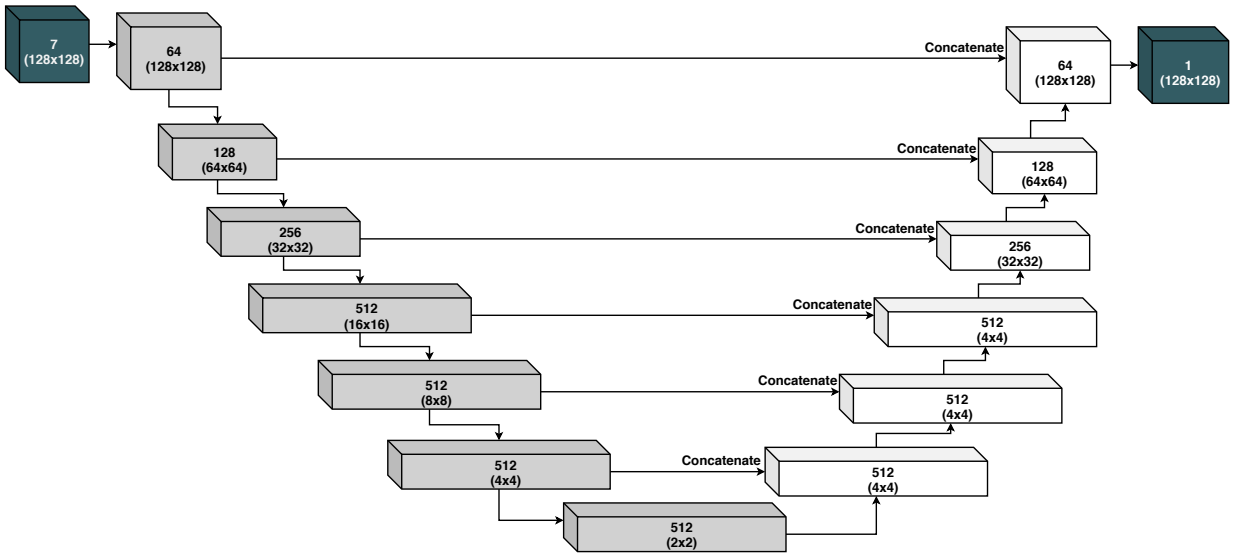


FIGURE 6.4: Schematic overview of the generator compartment of the used cGAN-network.

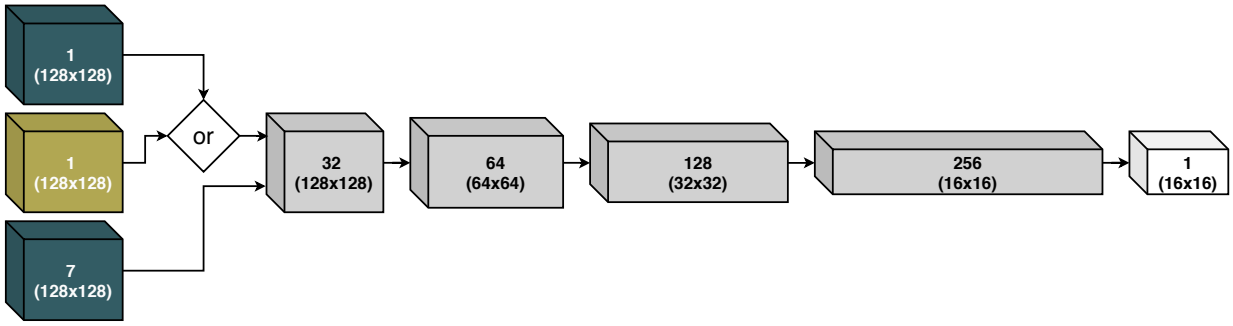


FIGURE 6.5: Schematic overview of the discriminator compartment of the used cGAN-network.

Discriminator

The second part of the cGAN is called the discriminator, and only has a function during training. Here, it's goal is to distinguish the synthetic CE-T1w scans from the real paired scans. The discriminator consists of four hidden layers with a configuration of 32-64-128-256 filters, with patches as output layer. All layers are composed of 3x3 convolutional kernes, and contain a LeakyRelu with a slope of 0.2 as activation layer (6.3 a). All activations were followed by batch normalization, except the first layer. A schematic overview is presented in figure 6.5.

Loss functions

The objective of this conditional GAN is given by equation 6.6:

$$\mathcal{L}_{cGAN}(Gen, Dis) = \mathbb{E}_{x,y}[log(Dis(x,y))] + \mathbb{E}_{x}[log(1 - Dis(x, Gen(x))] \,, \tag{6.6}$$

where *Gen* and *Dis* are the generator and discriminator respectively, x is the input image to be synthesized (DWI), y the reference image (CE-T1w). In equation 6.6, the goal of the generator is to minimize this function, and the discriminator to maximize it. To commit the generator to create synthetic images with high fidelity to the paired reference images, a L1 loss is added

between the synthetic and the true images:

$$\mathcal{L}_{L1}(Gen) = \mathbb{E}_{x,y}[||y - Gen(x)||_1] \,, \tag{6.7}$$

and is added to create the final objective (*) for the generator part:

$$Gen^* = \arg\min_{Gen}\max_{Dis} \mathcal{L}_{cGAN}(Gen, Dis) + \lambda\mathcal{L}_{L1}(Gen) \,, \tag{6.8}$$

where λ is a balancing parameter, and set to 100 in this work, similar to proposed in the original study displaying the conditional GAN *pix2pix* [168]. The single goal (*) of the discriminator is to minimize the errors in separating true from synthetic images, given the condition (DWI slice):

$$Dis^* = \arg\min_{Dis} \mathcal{L}_{cGAN}(Gen, Dis) \,. \tag{6.9}$$

Due to this $L_{cGAN}(Gen, Dis)$ objective, in addition to the regular $\mathcal{L}_{L1}(Gen)$, the generator is driven to synthesize real looking reconstructions.

6.1.4 Training

During training of the network, a batchsize of 8 slices was chosen. The input slices had a resolution of 128x128 pixels. The DWI slices were put in the layer channel of the batches. Shuffling was applied during on training data to present the data in random order each epoch. An Adam optimizer was used [170], a learning rate of $2.0 * 10^{-5}$ with parameters β_1, β_2 and $\hat{\epsilon}$ of $5.0 * 10^{-1}$, $9.99 * 10^{-1}$ and $1.0 * 10^{-7}$ respectively.

6.1.5 Model selection

During development of the model, various configurations of the cGAN were investigated. These experiments include multiple levels of depth for the U-Net, variations in batch normalization, dropout and data augmentation. It's effect on the performance was assessed quantitatively based on the validation set, but the final model for test experiments is decided on visual fidelity of the UF with paired CE-T1w scans. The influence of number of encoding generator layers is visualized in the Appendix, in section B.1.

6.1.6 Evaluation

The first part of the evaluation exists of visual examples of typical patients. Secondly, the data from the validation and test set is investigated. The reconstructions generated with the approach described in this chapter were assessed both quantitatively and qualitatively based on the data in the test set.

Quantitative evaluation

First, the distribution of pixel intensities in the synthetic reconstructions will be visualized in a boxplot for each ROI, together with the distributions of the pixel intensities in the true CE-T1w scans. Also, a normalized kernel density estimation will be graphed for pixel intensities in the synthetic CE-T1w scans, and a separate graph for the synthetic and true CE-T1w NPV and PV pixel intensities (Appendix). Also, the intensities of the synthetic and true CE-T1w scans will be analyzed on contrast differences between the PV and NPV on individual patient level. The fraction of patients where significantly different contrast is present will be reported in percentage. This analysis on individual patient level is only performed for NPV and PV ROIs, since those are deterministic for the ability to separate the ablated volume from viable tissue. In

addition for the quantitative evaluation, multiple objective measures are described, including the *mean absolute error* (MAE, equation 6.10), *mean squared error* (MSE, equation 6.11) and the *structural similarity index* (SSIM). The MAE is given by the following equation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \,, \tag{6.10}$$

where n is the sample size, y and $\hat{y}$ are the individual voxels of the true and synthetic CE-T1w scan respectively. The MAE measures the average magnitude of the differences between the predicted (synthetic) and true voxels, by weighing all individual errors equally. Thus, larger errors are not additionally penalized with MAE, opposite to other quadratic measures, such as the MSE:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \,. \tag{6.11}$$

Since the errors are squared before being averaged, the MSE results in high outcomes by relatively large errors compared to the MAE for example. A more recent approach to measure image fidelity is the *structural similarity* (SSIM) index [171]. The SSIM index takes local pixel aspects into account when assessing image fidelity. When postulating that images are highly structured and the pixels contain strong neighbor dependencies, the SSIM gives a more representing quantification of perceptual equality of the volumes than accustomed indices as the MAE for example. The basic idea of SSIM is comparing local patches a and b of two volumes at three similarity aspects: luminance (l), contrast (c) and structures (s). These similarities are combined to form a local SSIM:

$$SSIM = l(a,b)\cdot c(a,b)\cdot s(a,b) \,, \tag{6.12}$$

where $l(a,b)$, $c(a,b)$ and $s(a,b)$ are respectively:

$$l(a,b) = \frac{2\mu_a\mu_b + C_1}{\mu_a^2 + \mu_b^2 + C_1} \,, \tag{6.13}$$

$$c(a,b) = \frac{2\sigma_a\sigma_b + C_2}{\sigma_a^2 + \sigma_b^2 + C_2} \,, \tag{6.14}$$

$$s(a,b) = \frac{\sigma_{ab} + C_3}{\sigma_a\sigma_b + C_3} \,, \tag{6.15}$$

with μ_a and μ_b the local sample means of both the synthetic and true CE-T1w, σ_a and σ_b the local sample standard deviations, σ_{ab} the covariance of synthetic and true CE-T1w. C_1, C_2 and C_3 are small positive constants that stabilize each term, calculated by respectively $C_1 = (K_1, L)^2$, $C_2 = (K_2, L)^2$ and $C_3 = C_2/2$. L is the dynamic range of the pixel values, and K_1 and K_2 are two scalar constants, proposed respectively as 0.01 and 0.03 by the developers of SSIM [172].

The MAE and MSE two metrics are evaluated over the whole volume, and within the masks described in the ROI part from section 5.1.8. The SSIM is only measure for the whole volume. Before evaluating the volumes with the measures, both the true and synthesized volumes were rescaled to $[0, 1]$ intensity. This is done by rescaling the 1st and 99th percentile of all intensities to an intensity of 0 and 1 respectively. Secondly, a graph of probability density is presented for pixel-value within the separate NPV and PV regions post-HIFU. Clear separation between these distributions would represent high contrast between NPV and PV.

Qualitative evaluation (conducted by A. Zijlstra (2020), medicine student)

Qualitative analysis was conducted by means of questionnaires. Four radiologists with experience in MR-HIFU treatments from the Isala Hospital in Zwolle evaluated the synthetic and true CE-T1w scans unpaired in random order. The questionnaire included a section assessing the treatment progression of the MR-HIFU procedure. Whether the treatment was technically successful and finished based on the imaging data was asked in the first dichotomous question. In a following question, the observers were asked to estimate the NPV fraction, in percentages from 0-100%. In a separate section, the general aspects of the scan were assessed on a discrete 0-10 scale. This section included contrast quality, spatial resolution, SNR and overall quality of the scan. Zero indicated very poor quality, and 10 indicated perfect quality.

The two postprocedural MR-HIFU scans, i.e. synthetic and true CE-T1w, were shown in random order, and the observers assessed these scans with the questions above. The questionnaire also included more questions regarding specifics of the scans, those can be found in the internship report of A. Zijlstra (2020), the main findings are also listed here. When available, the observers had access to the screening synthetic and true CE-T1w scans for comparison to the postprocedural true or synthetic CE-T1w scan, during scoring. Also other scans from the screening or periprocedural MR-protocol where shown. The rationale of these questions was to evaluate whether the observers would assess the treatment progression differently on the synthetic CE-T1w scan compared to the true CE-T1w scan. If not, this indicates the synthetic CE-T1w scan could be safe and feasible to assess the MR-HIFU treatment progression.

6.1.7 Statistical analysis

In the quantitative analysis, D'Agostino's K-squared test [165, 166] with a significance level of 5% ($\alpha = 0.05$) is used to test if the measurements are taken from a normal distribution. Paired normal distributed samples are tested with a student T-test, while not normally distributed paired samples are tested with the Wilcoxon signed rank test. The mean contrast is tested on significant difference between the NPVs and PVs for the true and synthetic CE-T1w scans. On individual patient level, the contrast differences between intensities on true and synthetic CE-T1w scans within the PV and NPV ROI will be tested with a Mann-Whitney U-test. Also, the NPVs and PVs are tested between the scans, to detect whether they are sampled from the same distribution. An significance level of 5% is chosen for all tests.

For the qualitative evaluation, the dichotomous question regarding the technical success were analyzed with intra-observer agreement based on the responses of the true and synthetic CE-T1w scans. These results were reported by absolute agreement and 95% CIs. The estimations of NPV-ratio were tested on the significant differences between the true and synthetic CE-T1w scans, and the true ratios as measured on 3D on the true CE-T1w scans by a radiology resident. In case of normally distributed data a linear regression analysis was performed, and a Friedman's test for not-normally distributed data. The responses on a 0-10 scale assessing the general quality of the scans were analyzed with a paired T-test or a Wilcoxon's signed rank test after a normality test. On the general quality section, the tested aspects were considered sufficient when reached a score of ≥ 6 in more than 80% of cases.

6.2 Results

This section is set out with an overview of the patients from table 5.2 divided in three sets, and is provided in table 6.2. From the MaSS-II study population existing of 75 women, a DWI-scan was made from the 55 last treated patients. After exclusion of unusable sets, the following amount

of usable DWI datasets remained for testing and evaluation of the algorithm for the screening and post-HIFU scans:

set	screening (n)	post-HIFU (n)
training	19	21
validation	9	10
testing	15	19

TABLE 6.1: Total used DWI volumes in each subset for the development and testing of the DL-based approach, after exclusion of the unusable scans.

Visual examples of the final selected model are shown for three patients in figure 6.6 including DL-based reconstructions of synthetic CE-T1w scans, derived from DWI. These examples are from patients in the test set, and show three degrees of NPV-fractions.

Characteristic		Mean or count		
		D_{train}	D_{val}	D_{test}
Patients included		23	11	23
Age (years)		44.6 (± 6.05)	39.4 (± 8.58)	42.8 (± 5.99)
BMI (kg/m^2)		25.4 (± 3.13)	24.4 (± 4.43)	24.7 (± 3.51)
Abdominal fat layer (cm)		2.67 (± 1.33)	1.57 (± 0.94)	1.95 (± 0.97)
Uterine position	AVF	22 (95.7)	10 (90.1)	14 (60.9)
	RVF	1 (4.3)	1 (9.1)	4 (17.4)
	Upwards			5 (21.7)
Number of fibroids treated	1	19 (82.6)	7 (63.6)	15 (65.2)
	2	1 (4.35)	3 (27.3)	2 (8.70)
	3	2 (8.70)	1 (9.1)	4 (17.4)
	4	1 (4.35)		1 (4.35)
	5			1 (4.35)
Fibroids' location	Submucosal	10 (28.6)	2 (12.5)	10 (27.8)
	Intramural	7 (20.0)	5 (31.25)	13 (36.1)
	Subserosal	13 (37.1)	5 (31.25)	7 (19.4)
	Hybrid	5 (14.3)	4 (25.0)	6 (16.7)
Funaki class	1	3 (8.8)	1 (6.3)	7 (18.9)
	2	26 (76.5)	13 (81.2)	28 (75.7)
	3	5 (14.7)	3 (12.5)	4 (10.8)
Maximum diameter (cm)		5.28 (± 2.71)	8.25 (± 4.21)	4.70 (± 2.88)
Fibroid volume pre-HIFU (cm)		87.7 (± 138.9)	271.4 (± 321.1)	110.3 (± 185.3)
Non-Perfused Volume (%)		67.1 (± 33.9)	59.9 (± 30.1)	64.7 (± 34.8)

TABLE 6.2: Characteristics of the dataset visible in table 5.2, after division in the train, validation and test set.

6.2.1 Quantitative evaluation

The distributions of the pixel intensities in the post-HIFU synthetic CE-T1w reconstructions are summarized in a boxplot, visible in figure 6.7. In this plot, the means are compared within the different regions between the true and synthetic CE-T1w scans. The mean pixel intensities of the synthetic and true CE-T1w scans are for the *Full volume* $0.39(\pm 0.25)$ and $0.38(\pm 0.24)$ respectively. For the post-HIFU FV, the mean intensities for the synthetic and CE-T1w are $0.74(\pm 0.21)$ and $0.67(\pm 0.24)$ respectively. The PV had mean intensities $0.83(\pm 0.18)$ and $0.85(\pm 0.13)$ for the

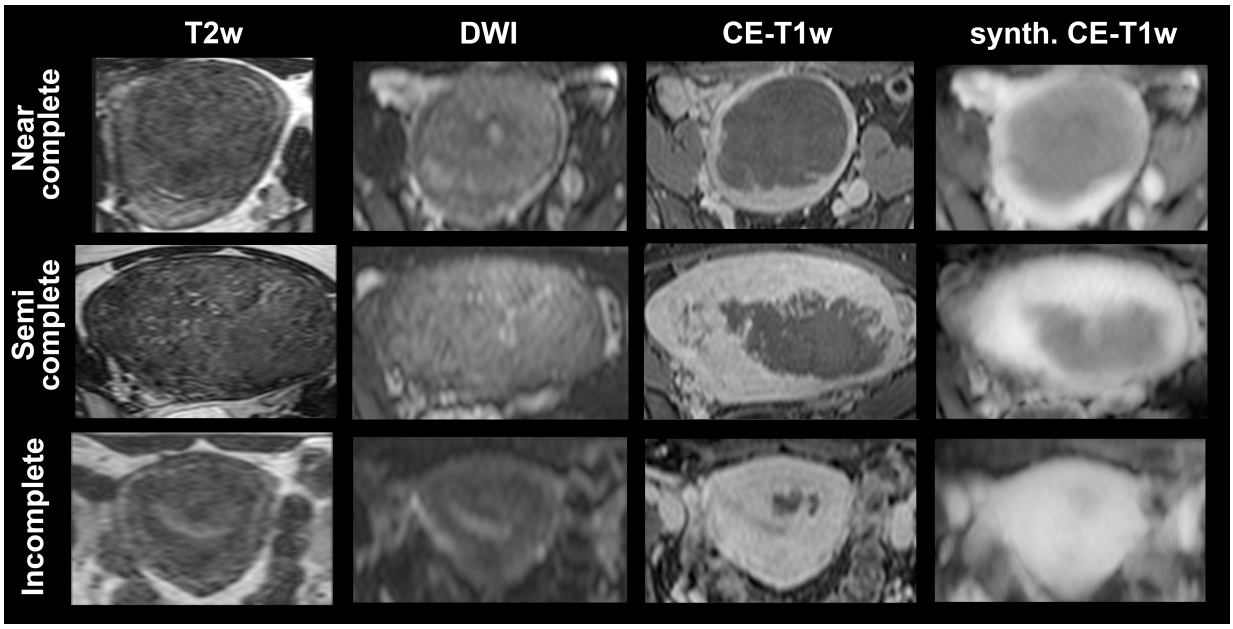


FIGURE 6.6: Overview of three UF reconstructions with the proposed DL-based approach. In this figure, patients with three levels of completeness of MR-HIFU ablation are shown, with the synthetic CE-T1w reconstructions in the last column, retrieved from the DWI scans in the second column. The reference true CE-T1w scans are depicted in the third column, and an anatomical T2w image is presented in the first column.

synthetic and true CE-T1w scans respectively. For the NPV, the synthetic and true pixel intensities were on average $0.64(\pm 0.20)$ and $0.47(\pm 0.14)$ respectively. In section B.2 the kernel density estimations are given (Appendix). On the level of individual patients, a statistically significant difference in pixel intensity between the NPV and PV ROI was found in 95% of patients for the synthetic CE-T1w scan, versus 100% for the true CE-T1w scans. In table 6.3 the data gathered with the metrics are listed, representing the mean outcome of the metric for all patients in the mentioned dataset.

	validation			test		
	MAE	MSE	SSIM	MAE	MSE	SSIM
Fibroid volume (FV)	0.189	0.175		0.227	0.218	
Non-perfused volume (NPV)	0.237	0.218		0.301	0.292	
Perfused volume (PV)	0.155	0.144		0.173	0.166	
Total volume	0.155	0.143	0.423	0.153	0.154	0.427

TABLE 6.3: Outcomes of metrics at the post-HIFU scans (true and synthetic volumes were normalized to $[0, 1]$ intensity in different anatomical regions). These outcomes describe the similarities in pixel intensities between the true and synthetic CE-T1w scans. For the MAE and MSE metrics, an outcome close to zero denote good resemblance of both scans, while an SSIM of 1 resembles equal scans.

Statistical analysis

The average of the mean intensities for all test patients in the NPV and PV regions for the true CE-T1w scans are $0.48(\pm 0.088)$ and $0.82(\pm 0.10)$ respectively. For the synthetic scans, the averages for NPV and PV intensities are $0.75(\pm 0.15)$ and $0.81(\pm 0.14)$ respectively. From the

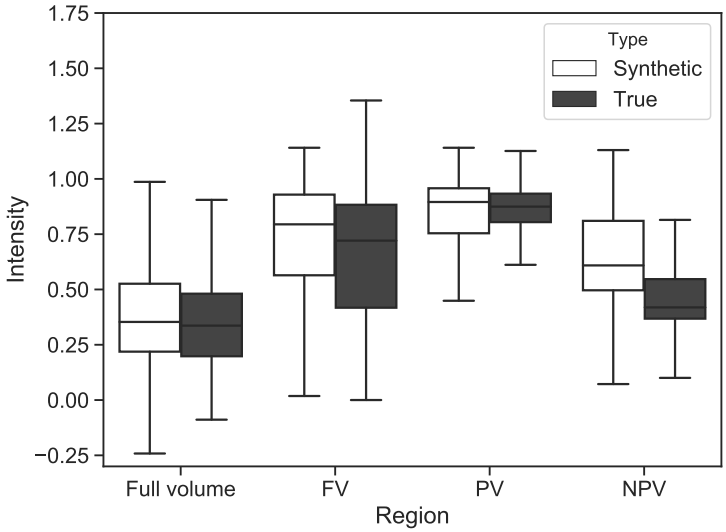


FIGURE 6.7: Distributions of voxel intensities of the true (dark) and synthetic (light) CE-T1w scans within different anatomical regions. The goal of the DL-training was to match the distributions of the true and synthetic CE-T1w scans properly. The clinically relevant contrast is mainly resembled by the difference in intensity of the *perfused volume* (PV) and the *non-perfused volume* (NPV). As can be seen, the contrast between those regions is larger at the true CE-T1w scans, but the intensities at the synthetic CE-T1w scans also differ significantly between these regions (as shown in this work). Another appearing phenomenon is the larger spread in the NPV intensities for the synthetic CE-T1w scans. This indicates that the pixel intensities of non-perfused voxels are more often high for the synthetic CE-T1w scans in relation to the true CE-T1w scans.

D'Agostino and Pearson's normaltest, the means in both true and synthetic NPV were normally distributed ($p = 0.25$ and $p = 0.74$ respectively). The PVs from the true and synthetic CE-T1w scans were both not normally distributed with a p-value of $7.43 * 10^{-4}$ and $5.18 * 10^{-3}$ respectively. Therefore, the Wilcoxon signed-rank test was used to investigate the eventual difference between the distributions when not normally distributed, and the paired t-test was used otherwise.

The distributions of the mean PV and NPV intensities were significantly different in the test set for both the true and synthetic CE-T1w scans: $p = 8.86 * 10^{-5}$ (true) and $p = 8.97 * 10^{-3}$ (synthetic). The PVs for both true and synthetic scans did not differ significantly with a p-value of 0.85, while the distributions of the NPVs were statistically different ($p = 2.31 * 10^{-7}$).

6.2.2 Qualitative evaluation

The results of the section of the qualitative evaluation that asked whether treatments were technically successful based on the true and synthetic CE-T1w scans are listed in table 6.4. The responses on the cases with a complete scan protocol have been separately reported. For the cases from the full protocol, absolute agreement between assessment based on the true and synthetic CE-T1w scans was achieved in 50/60 cases (83%). The majority of the disagreement of the cases were caused by false-negatives on the synthetic CE-T1w scans with a false-negative rate of 18%, while the false-positive rate was 9%. True-positive and true negative rates were respectively 82% and 91%. The absolute, positive and negative agreement of cases with the total scan

post-procedural assessment of technical success					
	on synthetic CE-T1w	**on true CE-T1w**			
		All patients		**Full protocol**	
		Yes	No	Yes	No
Observer 1	Yes	9	1	8	0
	No	4	5	3	4
Observer 2	Yes	12	2	11	0
	No	4	1	3	1
Observer 3	Yes	11	3	10	1
	No	3	2	2	2
Observer 4	Yes	13	1	11	0
	No	1	4	1	3
All observers	Yes	45	7	40	1
	No	12	12	9	10

TABLE 6.4: Agreement between the answers of the observers when asked when the treatment was technically successful based on the synthetic or true CE-T1w scans. During the test, the observers assessed the technical outcome of the MR-HIFU procedure based on the synthetic and true CE-T1w scans in random order. The cases with an incomplete scan protocol have been separated, and are listed in the column 'Full protocol'.

	$n(\%)$ subjective assessments ≥ 6	
	True CE-T1w	Synthetic CE-T1w
Contrast	76/76 (100)	70/76 (92.1)
Spatial resolution	75/76 (98.7)	70/76 (92.1)
SNR	74/76 (97.4)	71/76 (93.4)
Total quality	76/76 (100)	70/76 (97.4)

TABLE 6.5: General aspects of the true and synthetic CE-T1w scans, with combined responses of the four observers, denoting the fraction that scored sufficiently ($\geq$ 6). More than 80% was considered sufficient here, and both scans reached that percentage on all aspects.

protocol are listed in Appendix section B.3, including the 95%-CIs.

For the NPV fraction estimation, no statistical differences were found in the estimations based on the synthetic and true CE-T1w scans ($\alpha = 5\%$). Also, no statistical difference when comparing the estimations of both scans with the true NPV-ratios, as measured on the true CE-T1w scan in 3D. The p-values of this multi-group analysis for the first, second, third and fourth observer were 0.052, 0.437, 0.115 and 0.951 respectively, implicating no difference between the three distributions (i.e. NPV ratio estimated on synthetic and true CE-T1w scans, and ground-truth ratios measured on true CE-T1w scan in 3D). For the exact Friedman outcomes, including percentile descriptives and ranks, see Appendix section B.4.

In the qualitative analysis section assessing the general quality of the scan, all aspects (contrast, spatial resolution, SNR, and total quality) were significantly lower on the synthetic CE-T1w scan than the true CE-T1w scan (p<0.0005). Both scans achieved sufficient scores ($\geq$ 6 on more than 80%) on all measured aspects, as listed in table 6.5. This implicates despite the synthetic CE-T1w scans having inferior image quality relative to the true CE-T1w scans, their image quality can be considered sufficient.

Chapter 7

Discussion

7.1 General discussion

In this work, two methods were revealed for contrast agent-free visualization of the NPV after or during MR-HIFU treatment of UF. With the DWI IVIM approach described in chapter 5, the f-parameter showed statistically significant contrast differences between the NPV and surrounding perfused tissue, on a post-procedural DWI scan. Also with the second approach, based on DL, statistically significant contrast differences were found quantitatively between the PV and NPV. Subjective analysis with radiologists established the image fidelity of the synthetic CE-T1w scans relative to paired true CE-T1w scans. NPV ratio estimations on the true and synthetic CE-T1w scans were not statistically significant different.

Based on these findings, the shown methods in this work are a viable Gd-free alternative to CE-T1w, in order to assess the NPV after MR-HIFU treatment of UF. An implication of this is the possibility that procedure times and NPV-rates can be improved for MR-HIFU treatments, when contrast agents are no longer a necessity to evaluate technical treatment effects. Periprocedural DWI-scans can be translated to f-maps or synthetic CE-T1w scans to aid the operator in visualization of treatment effect, and determining the progress of ablation. Hence, procedure times can be shortened by avoiding unnecessary prolongation of ablations, and NPV-rates can be improved by circumventing premature stopping of ablations.

7.2 IVIM approach

It was shown here quantitatively that the NPV can be distinguished from the post-HIFU PV and the FV by using DWI IVIM. Parameter maps were acquired by fitting the IVIM equation (equation 5.6) to the DWI datasets with a least-squares method. A significant difference was found in f between the NPV and the post-HIFU PV and FV during screening. Also, the D parameter deviated significantly within the post-HIFU PV and the FV during screening. These findings suggest that the f-maps could be used to evaluate the treatment progress without using Gd-based contrast agents.

The f-parameter decreases significantly during ablation of tissue with MR-HIFU, as can be seen in the boxplot in figure 5.5 (b). It was shown that in 89% of patients the average intensity of the NPV and PV on the f-maps differed significantly. Since in all these 89% patients a decrease in f-value was found, the f-maps show an consistent local decline in intensity after MR-HIFU ablation, and are therefore an interesting solution for Gd-free visualization of the MR-HIFU treatment effect. The location, size and geometry of the NPV and unablated PV were defined on the paired CE-T1w scan, by drawing ROIs. The decrease of f reflects the reduction of the fast attenuation component in the DWI signal relative to the slow, i.e. diffusion component of

the signal. This component describes the incoherent motion of water molecules in the capillary network [149], and is hypothetically related to the degree of microcirculation, or perfusion of tissue. This means when f is reduced, the volume fraction of microcirculation is attenuated in relation to the diffusion component.

A possible explanation of the decrease in f after MR-HIFU treatment might be that after successful ablation necrosis occurs, entailing decrease in activity of capillaries and a reduction in flowing blood volume. This hypothesis is consistent with previous research, reporting coagulation necrosis and devascularization after HIFU ablation [173, 174].

Also, the mean D-parameter showed a significant decline from the total FV during screening, and the unablated PV after MR-HIFU treatment, as visible in the boxplot in figure 5.5 (a). This means that the diffusion was more restricted in the PV and diffusion rate had lower velocity. This phenomenon may be due to the intracellular uptake of fluid, increasing the volume of intracellular water. Fluid in the intracellular compartment is more restricted than in the extracellular interstitium [175]. It can therefore be assumed that at the boundaries of NPV, diffusion gets more restricted, and intercellular fluid increases. This might caused by partially ablated cells, that are not completely necrotized. An intriguing follow-up question would be to ask whether longterm outcomes are related to the described effects visible on the D-map. It can be hypothesized that in particular the locations with decreased diffusivity after MR-HIFU ablations will recover and remain viable in a relatively short period of time.

In terms of diffusion values inside the NPV compared to the unablated PV, it can be seen in figure 5.5 that there is a slight increase in diffusion value. Although this effect is not statistically significant for the total dataset ($p < 0.05$), this rise might reflect the increased extracellular fluid due to necrosis of cells, releasing the intracellular fluid in the surrounding interstitium. It should be noted however, that the accordance between the CE-T1w and the DWI sequences is not perfect. This can partly be explained by the contrast solution diffusing in the NPV, entailing an underestimation of the true NPV. If focusing on the individual patient level, it was established that the average D-value within the NPV was significantly higher in 58% of cases compared to the PV, and lower in 31% of patients. This suggests that the diffusivity in tissue changes relatively common after MR-HIFU ablation (in 89% of total patients), but can either increase or decrease compared to the viable surrounding UF tissue. It is likely that this phenomenon largely depends on UF characteristics, and further research is needed to elaborate on this finding to explain the cause.

These effects of D and f were also visualized for a single patient with near-complete ablation if figure 5.4. It can be seen that despite the noise, a clear low-intensity ring is visible in the D-map, on and around the borders of the NPV. For the f-map, the hypo-intense region is easily distinguishable in comparison to the screening f-map, and correlates to the NPV showed on the CE-T1w scan.

Comparison to literature

In 2005, the first DWI paper regarding assessment of MR-HIFU treatment effects in UFs was published by Jacobs et al. (2005) [176]. The authors used a mono-exponential model (equation 2.11) to describe DWI data with three b-values (0, 500 and 1000 s/mm^2), and demonstrated feasibility of DWI and ADC mapping for identification and monitoring of the NPV after MR-HIFU.

Bi-exponential models (equation 2.13) as proposed in this work are not known in assessment of treatment effects after MR-HIFU ablation of UFs. However, it has already been suggested

by Ikink et al. (2014) [177] that low b-values are useful in emphasizing perfusion effects after MR-HIFU. This encourages the hypothesis that the perfusion effects influences the DWI values in the low b-values, and could be extracted by a bi-exponential model as proposed here.

Although IVIM has not yet been performed as strategy to assess MR-HIFU treatment effects, it has been used in evaluation of UAE in the study from Cao et al. (2017) [178]. The authors described twelve patients with a DWI sequence with b-values 0, 50, 100, 200, 500, 800, 1000, 1200 s/mm^2 at baseline and six months follow-up, and found also a significant decrease in f and an increased ADC value. The increase in D-value and decrease in f-value found here, as previously stated, confirms findings in previous research commenting on a high-signal ring on DWI, collocating the NPV on CE-T1w [179].

Another study from Guo et al. (2015) [180] used 16 b-values between 0 and 800 s/mm^2 to perform IVIM analysis after radiofrequency ablation of tumors in 10 rabbits. They demonstrated that IVIM that changes in D, D^* and f can be utilized to predict tumor response after radiofrequency ablation, but they found no evident link between the parameters and dynamic CE-MRI. These findings might raise intriguing questions regarding the nature and extent of IVIM analysis to describe the perfusion of tissue. However, it should be noted that the perfusion characteristics and blood flow dynamics in a tumor are likely to differ by large amount from UFs.

The relationship between IVIM parameters and CE-MRI has been studied extensively previously, and resemblance between IVIM and (dynamic) CE-MRI seems to be highly application dependent since it still remains a much debated question [180–184]. Therefore, the possible link between CE-MRI and IVIM may be limited by the specifics of the application it is investigated in.

7.3 Deep learning-based approach

The strategy described in chapter 6 was designed to determine the effectiveness of a DL-based method in translation of DWI to CE-T1w. The results of this chapter show that the proposed DL-based network (cGAN) was feasible as method for synthetic CE-T1w generation. This was supported by both the quantitative and the qualitative evaluation performed in this work.

With the quantitative evaluation is was established that significant differences in intensity was achieved within patients in the NPV relative to the unablated PV, similar to the f-parameter from the IVIM analysis (see boxplot for intensity differences 6.7). The intensities of the PVs on the true and synthetic CE-T1ws were not statistically different, while the NPVs were. The mean NPV intensities on the synthetic CE-T1w were higher than on the true CE-T1w scans. An implication of this might be that parts of non-perfused areas could resemble viable perfused tissue on synthetic CE-T1w scans. This inconsistency could be attributed to the choice of network, that was partly based on the least ratio of false NPV appearance on the synthetic CE-T1w on scans from the validation set, see e.g. section B.1. It is possible, therefore, that another model producing synthetic CE-T1ws with more contrast between NPV and PV do not result in statistical differences between synthetic and true NPVs.

It can be seen from the graph in figure 6.7 that the distribution of pixel intensities after an MR-HIFU treatment is steeper for true CE-T1w scans than synthetic CE-T1w scans. One of the factors explaining this increased spread at the synthetic scans are the imperfect registration between the DWI and CE-T1w scans. This leads to geometrical differences of the synthetic and true CE-T1ws, leading to interference of different tissue types within in the drawn ROIs, based

on the true CE-T1w scans.

Important findings from the qualitative evaluation include the result that radiologists decided equally on the synthetic and true CE-T1w scan whether the treatment was technically successful in 83% of the cases. It was also found that both the NPV ratio estimations on the synthetic and true CE-T1w scans did not differ significantly from the true measured NPV ratios. This means that on both scans, the ratio between the NPV and the FV can be adequately estimated by the radiologists. The last main qualitative finding was the difference in general quality between the synthetic and true CE-T1w scan. Although the quality on the measured aspects of the synthetic CE-T1w was significantly lower compared to the true CE-T1w scan, the quality was still sufficient.

An noteworthy finding was the difference in qualitative performance of the radiologists on the reconstructed synthetic scans when the full scan protocol was available, compared to protocols with missing data (mostly screening scans). We found that assessment based on the synthetic CE-T1w scans of the technical success increased in agreement with true CE-T1w scans when the full protocol was available (75% improved to 83%). This emphasizes the importance of a (synthetic) scan acquired during screening for comparison, when assessing a post-HIFU scan.

7.3.1 Comparison to literature

Up to now, very little attention have been paid to DL-based methods for replicating gadolinium-based contrast on MRI. Moreover, no previous study has investigated synthetic CE-MRI as tool for periprocedural MR-HIFU treatment evaluation. What has been investigated to a broader extent is the DL-based dose reduction for Gd-based contrast agents [185].

DL-based production of Gd-contrast was recently proposed by Sun et al. (2020) [186] in brains of 49 mice, based on noncontrast T2w scans. Another study that approaches the efforts from this work in synthetic Gd contrast generation the closest was conducted by Kleesiek et al. (2019) [187] They proposed a method for predicting contrast enhancement from noncontrast enhanced multiparametric brain MRI including a DWI sequence, also using a DL-approach. However, they only used two different b-values, i.e. 0 and 1200 mm^2/s. However, it has been established that information regarding perfusion of tissue is present in low b-values [177]. For quantitative evaluation, the authors reported PSNR values ranging between 30 and 20, and an SSIM between 0.74 and 0.88.

7.4 Limitations

Since the nature of this study was retrospective, collection of data have not been conducted with IVIM and DL-based methods in mind. In hindsight, consistency of scanned anatomical regions and unvaried acquisition of screening scans would have likely increased training of the DL-network and statistical power of the analysis conducted here. Nevertheless, the acquisition of multi-parametric data have been performed aiming at a comparable goal as stated in this thesis (section 4).

Also, the generalisability of these results are subject to limitations. The results shown in this work are based on MR-data collected from a single center. It has been established that acquisition of MR-scans are widely vendor and scanner specific [188–190] and, therefore, generalisability of results presented here is likely impeded by this factor. Moreover, the spread in MR-HIFU

treatment results is not homogeneous, as during the inclusion of patients a learning curve was present, resulting in on average higher NPV ratios in the later included patients.

Regarding the analysis in chapter 6, an issue that was not addressed in this study was the intraobserver consistency within the true CE-T1w scans. An implication of this is that the inherent intraobserver consistency remains concealed, while interfering with the consistency measurements between true and synthetic CE-T1w. Besides, a higher number of experts could lead to better insights in the true perceived similarity between true and synthetic CE-T1w. Also, the observers that participated in this study were from the a single center, where the data was also collected. General performance of the synthetic CE-T1w scan, relative to assessment by a radiologists, would be better evaluated if the background of the observers had more variety.

For the development and training of the DL-network, optimal hyperparameters have not been searched with a systematic grid-search, but had been based on literature and prior experience. For this specific application, it is plausible that a optimum in hyperparameters differs from the hyperparameters described in this work. Taking into consideration that the search space of the optimal combination of hyperparameters is large, ensuring the best set of parameters is a difficult task.

7.5 Future recommendations

A natural progression of this work is to analyze the effectiveness of both IVIM and the DL-based method in a prospective manner. Further work needs to be done whether these methods have a significant added value in clinical practice, and indeed lead to improved treatment outcomes and shorter treatment times.

An interesting suggestion for future efforts is DL-based IVIM fitting, replacing the conventional least squares fit. With a successful implementation, as shown feasible by Barbieri et al. (2020) [191], could lead to precise, accurate and in particular fast IVIM model fitting to DWI data. This technique has been exploratory investigated by the author of this work. The preliminary results look promising, however following systematic investigations are required before publication.

Besides treatment evaluation, the extraction of IVIM parameters from the DWI data could also have potential to deliver insights in prediction during screening and follow-up of MR-HIFU. For screening, IVIM parameters might aid the process of predicting treatment outcomes, and help in patient selection for the MR-HIFU treatment. For follow-up, IVIM parameters could hypothetically serve as predictor for long term outcomes, as discussed earlier. Future work should further elaborate the relation between the IVIM parameters and these topics.

With the availability of a tool to assess treatment outcomes periprocedurally, continued efforts could eventually lead to efficient application of targeted vessel occlusion, potentially entailing short treatment times [177]. With such a tool, these hypothesized effects can be further investigated and eventually utilized to create an improved non-invasive treatment option for patients suffering from UFs.

To further investigate and exploit the phenomenon of finding a larger periprocedural NPV than expected based on the placement of the transducers focal point, an improved screening MR-protocol would be helpful. At this moment, the enlarged NPV occurs sporadically and is hard to anticipate. The main reason for this is the usual ignorance regarding the exact location

of the nourishing vessel(s) of the UFs. To be able to exploit the enlarged occurrence of NPV, hypothetically caused by vessel occlusion, the location of the incoming vessels of the UF should be localized during screening. This requires improved visualization of the UFs vasculature during screening, or directly preprocedurally. To achieve more insights in the circulation of an UF, MR angiography or US-based power doppler could be embedded in the screening protocol.

In addition, to gain progress in the discussion around the exact mechanism behind MR-HIFU induced vessel occlusion or ablation, a pilot study investigating the histopathology of MR-HIFU sonications on vessels could be helpful. A proposal of a study design covering this topic would be to include women with UFs undergoing hysterectomy in the near future. Before surgery, these patients will receive MR-HIFU treatment with the targeted vessel occlusion strategy. After subsequent hysterectomy, histopathology could examine the resected UF from surgery and describe the effects of the MR-HIFU ablation.

Lastly, future research should be undertaken to not only improve modeling, but also acquisition of improved DWI-data. This work made use of a range of b-values between 0 and 800 mm^2/s, but a larger number of, in special, low b-values are likely to reduce noise, and improve model outcomes. Implications of this could be a higher spatial resolution and lower uncertainties with predictions. Moreover, the clinical practice could benefit when faster acquisition leads to shorter treatment times. In addition, the used DWI-data was single-center and single-vendor, and should ideally be multi-center and multi-vendor, in particular for DL applications [192]. Future work around this topic should therefore aim at wider data acquisition.

7.6 Conclusion

7.6.1 General conclusion

Many women suffer from uterine fibroids at some moment in their life. These benign tumors can lead to complaints that require treatment to achieve symptom relieve. Although a non-invasive uterus-sparing treatment option is available, called MR-HIFU, the invalidating surgical options are more commonly opted for. One of the reasons for this is the limited availability of MR-HIFU. This is due to long procedure times, as MR-HIFU treatments of uterine fibroids generally take a long time. This is partly caused by the inability of the operating radiologists to view the treatment progression during the MR-HIFU procedure. Currently, the treatment progress can only be assessed after injection of a contrast-agent. A contrast-agent can unfortunately only be injected once, when the treatment has already been finished.

Thus, procedure times can possibly be shortened when no contrast-agent is needed for visualization of the treatment progression. This thesis attempted to offer a solution for this problem. Two methods are proposed in this work to achieve this goal of visualization of the treatment effect without needing a contrast-agent. By analyzing a specific MRI-scan with a mathematical equation and an artificial intelligence model, it has been shown that these proposals are adequate for visualization of the MR-HIFU treatment progression. The efforts in this work contribute to the goal of making MR-HIFU more reliable and available as treatment option for symptomatic uterine fibroids. With the accomplishments from this work, more women will eventually be able to benefit from the non-invasive attractiveness of MR-HIFU as treatment option for symptomatic uterine fibroids.

7.6.2 Detailed conclusion

This work has revealed two effective methods for Gd-free visualization of the NPV after MR-HIFU treatment of UF by using DWI, enabling peri- and intraprocedural treatment evaluation. It has been shown quantitatively that both the IVIM and DL-based strategies result in significantly different pixel intensities between the unablated PV and the ablated NPV, suggesting sufficient contrast for discrimination between them. Qualitative evaluation of the DL-based approach disclosed an agreement of 83% between clinical decisions on the reference and generated CE-T1w scans. NPV ratio estimations could be adequately performed on the synthetic and true CE-T1w scans, with no significant variation from the ground-truth NPV ratios. The findings and proposed methods in this work can lead to improved outcomes of MR-HIFU treatments. Moreover, with feasible and safe Gd-free intraprocedural visualization of the MR-HIFU treatment progression, new avenues are opened for testing treatment techniques that possibly shorten procedure times. These implications benefit the patient suffering from UF, by making a non-invasive treatment option more reliable and available, leading to more efficient healthcare.

Appendix A

Appendix IVIM

A.1 Trust-region optimization

The TRF method is used to find the minimum of an objective function (in this case equation 5.8). Therefore, the sum of the squared residuals $S(\beta)$ is minimized by finding the optimal minimizer β containing the model parameters. $S(\beta)$ is also called the objective function, and the general goal of TRF is to find a local solution for minimization of $S(\beta)$.

In order to illustrate the basic concepts of TRF, suppose an example where the objective function $S(\beta)$ has only two parameters β^1 and β^2. The objective function relative to these parameters may look like the contour map in figure A.1.

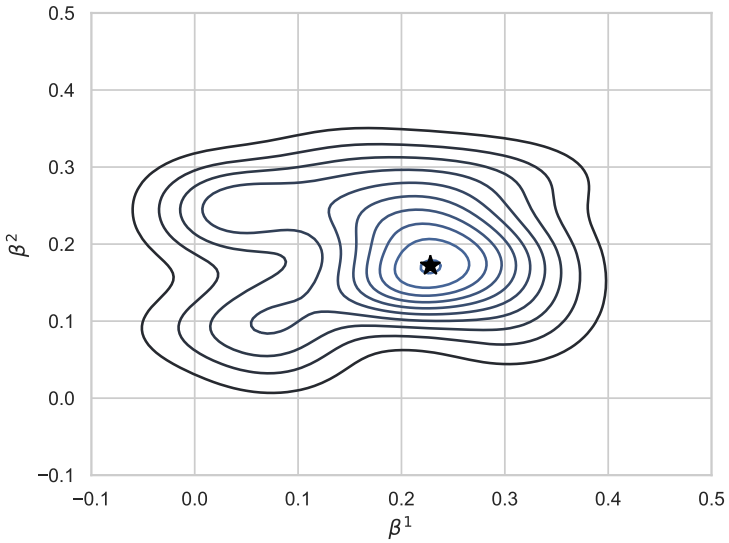


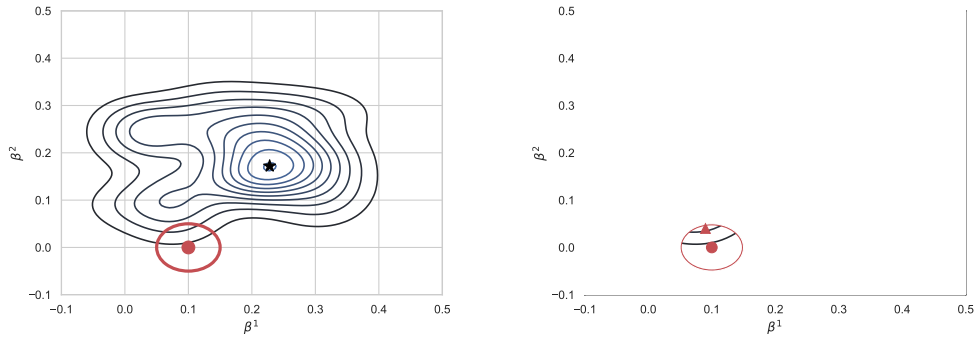
FIGURE A.1: Contour map of objective function $S(\beta)$ relative to β_1 and β_2. A local minimum is located at the star-symbol.

With the TRF algorithm, the goal is to find the (preferably global) minimum at the star-symbol in figure A.1, i.e. finding the β parameters that minimize the sum of the squared residuals. The TRF method is an iterative algorithm, that starts at $t = 0$ by defining an initial estimate of the parameters: β_0. At each next iteration, new values for β are searched that are closer to the local minimum. To achieve this, a model $\widetilde{S}_t$ of the objective function S is calculated:

$$\widetilde{S}_t(\beta_t + s) = S(\beta_t) + g_t^T s + \frac{1}{2} s^T B_k s \,, \tag{A.1}$$

where $\widetilde{S}_t(\beta_t + s)$ is the quadratic approximation of S based on the second degree Taylor expansion around β_t. The gradient of $S(\beta)$ at β_t is defined by g_t and B_k is an approximation of the second order derivative of $S(\beta)$ at β_t. With $\widetilde{S}_t(\beta_t)$ the function of $S(\beta)$ is estimated around β_t. In general the relation between the objective function and the Taylor expansion loosens at a greater distance to the expansion point (β_t). During the TRF iterations this model $\widetilde{S}_t(\beta_t)$ is trusted in an area with radius λ_t to the current point of β_t.

Specifically, at the start of the iterations β_0 is defined, for example at $\{0.1, 0\}$ in the graph of figure A.1, with a trusted-region with radius 0.05, as depicted in figure A.2.



(a) (b)

FIGURE A.2: (a) Contour map of $S(\beta)$ with β_0 (red dot) and the trusted-region λ_0 (red circle). (b) Contour map of $\widetilde{S}_0(\beta_0)$ inside the trusted-region λ_0 at $t = 0$.

After defining β_0, a step c_t with a 'sufficient' negative gradient is searched within the trusted-region of $\widetilde{S}_t$. This is done by finding the *generalized Cauchy point* of the model, that is the first local minimizer of equation A.1 [193]. When a trial step is found, the ratio of the objective function and model at this trial point is calculated with the following equation:

$$\rho_t = \frac{S(\beta_t) - S(\beta_t + c_t)}{\widetilde{S}_t(\beta_t) - \widetilde{S}_t(\beta_t + c_t)}, \tag{A.2}$$

and then 1 of three steps is performed based on the outcome of ρ_t:

1. ρ_t is larger or equal than a given threshold: the new trial point is accepted and $\beta_{t+1} = \beta_t + c_t$,
2. ρ_t is smaller than the threshold: the new trial point is declined and $\beta_{t+1} = \beta_t$.

Following the above situations: the trusted-region's radius λ_{t+1} is increased in case 1, when ρ_t exceeds the preset threshold. Accordingly, in the second case the trusted-region is decreased by a preset amount when $\widetilde{S}_t(\beta_t + c)$ did not bear enough relationship to $S(\beta_t + c)$.

Suppose in the example in figure A.2, new trial point at the triangle symbol results in a large enough ρ_0, the point is accepted and the $\beta_1 = \beta_0 + c_0$, and the trusted-region λ_1 is increased relative to λ_0. The result of these steps is visualized in figure A.3 (a). The process of finding a new trial point, acceptance, moving to the next iteration $t = 2$ and subsequent enlargement of the trusted-region radius to ρ_2 results in figure A.3 (b).

Now, the trusted-region covers a large part of the search space of β, but this map of $\widetilde{S}_2(\beta_2)$ does not bear sufficient relation to the true map of $S(\beta)$, as can be seen when comparing A.3 (b) with A.1. This time, the new found trial point at the triangle symbol is rejected as ρ_2 does not

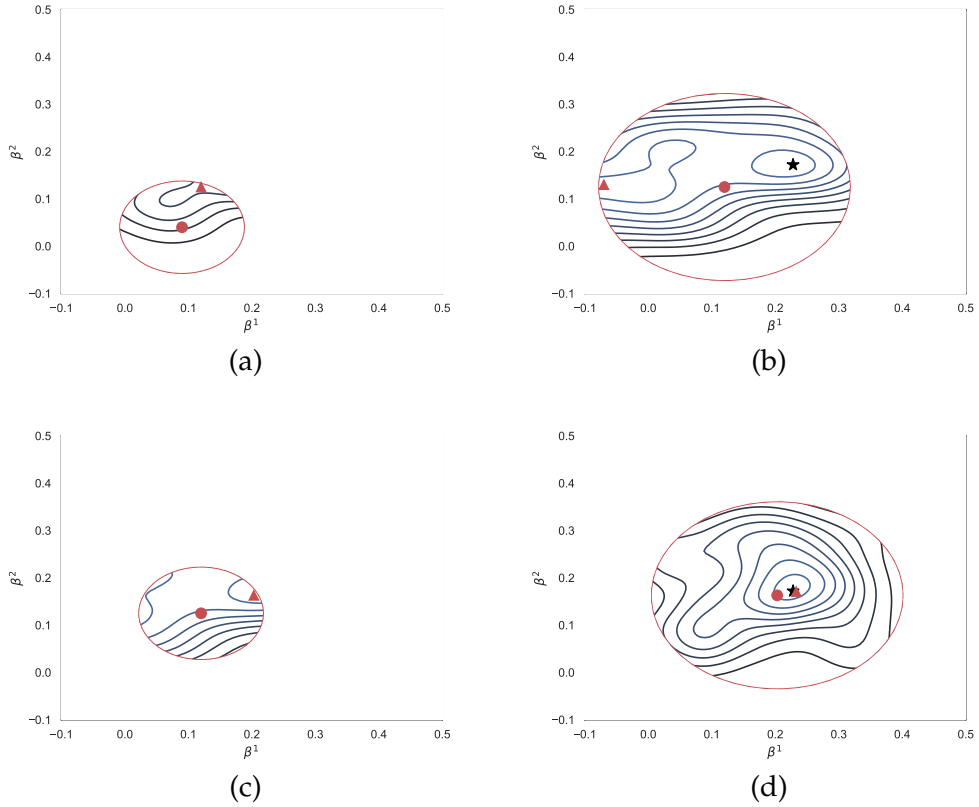


FIGURE A.3: Contour maps of $\widetilde{S}_t(\beta_t)$ with the red dot denoting β, the red circle the trusted region with radius λ_t and the red triangle symbol representing the trial point $\beta_t + c_t$, for $t \in \{1,2,3,4\}$. (a) Contour map of $\widetilde{S}_1(\beta_1)$ inside the trusted-region λ_1 at $t = 1$ after acceptance of c_0. (b) Contour map $\widetilde{S}_2(\beta_2)$ after acceptance of c_1, with the red triangle symbol at trial point $\beta_2 + c_2$ with low relation to $S(\beta_2 + c_2)$. (c) Contour map of $\widetilde{S}_3(\beta_3)$ after declining c_2, at this iteration with a smaller trusted-region, c_3 does bear sufficient relationship to $S(\beta_3 + c_3)$. (d) Contour map of $\widetilde{S}_4(\beta_4)$ after accepting c_3, the trusted-region is enlarged again and the minimum (star symbol) is closely approached by the next step c_4.

pass the preset threshold. Therefore case 2 is initiated: $\beta_3 = \beta_2$ and $\lambda_3 < \lambda_2$, resulting in figure A.3 (c). At this iteration, a trial point is searched in a smaller trusted-region, and ρ_3 is large enough. The trial point is accepted at the trusted-region enlarged again, according case 1, illustrated in figure A.3 (d). At this moment, the new trial point is in the vicinity of the true local minimum at $S(\beta)$, and once the distances between subsequent updated trial points are small enough, the algorithm can be stopped and the local minimum can be estimated.

A.2 Scatter plot IVIM for relation D with f

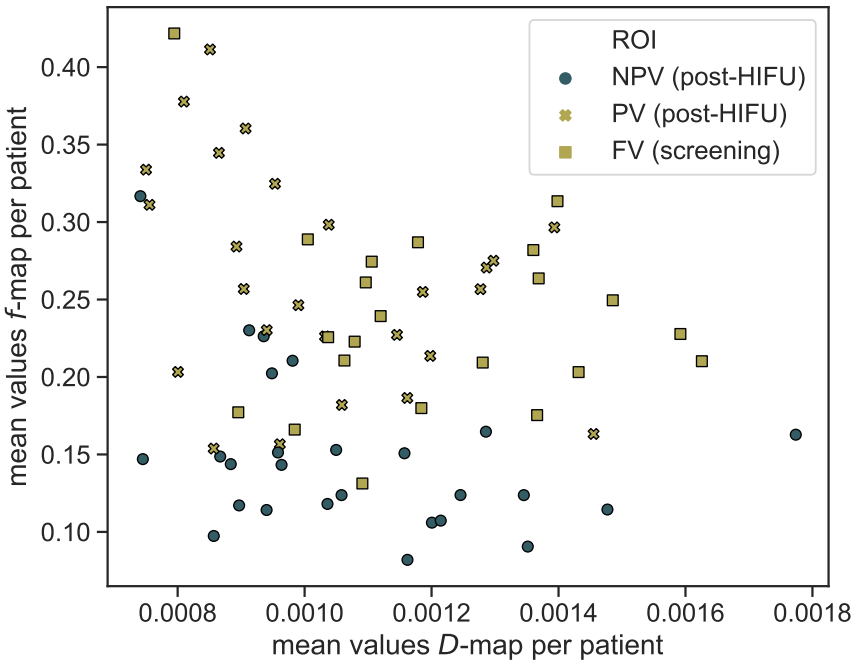


FIGURE A.4: Scatter plot of the mean D versus the mean f parameters found with the IVIM-fit in the three different regions (i.e. screening FV, post-HIFU PV and post-HIFU NPV). Each marker represents the mean value in a patient. With this graph, the relation between D and f in these patients is visualized.

Appendix B

Appendix DL-based

B.1 Generator depth

Figure B.1.

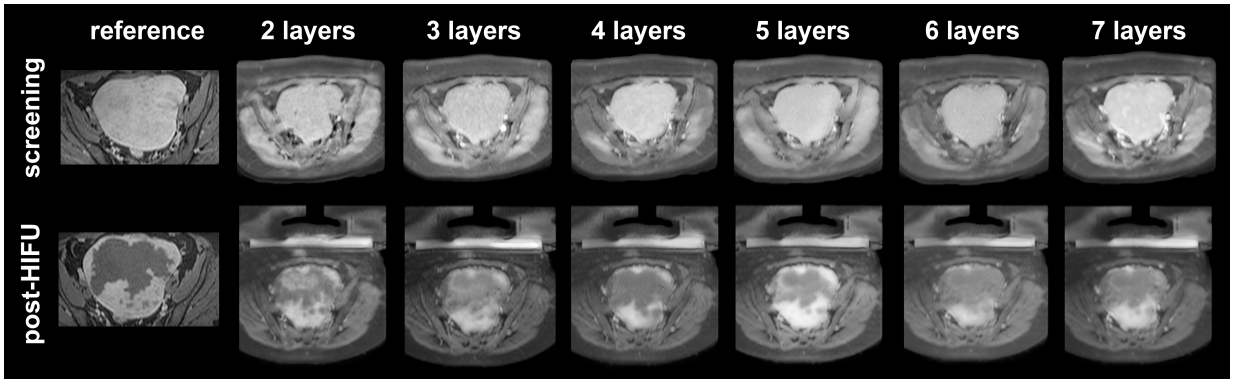


FIGURE B.1: Influence on DL generator depth on predicted volume, see also figure 6.4 for the generator architecture. The model investigated in this work consisted of seven encoding layers. The encoding generator layers have been varied between two and seven layers, and are visualized in this figure for axial reconstructions of the female pelvic region including an UF. Although no clear trend can be extracted, it can be derived that the clarity of the NPV delineation is not depending on the depth of the generator alone, as shallow (two and three layers) and deep (seven layers) networks seem to delineate the NPV relatively clear compared to the network with five encoding layers. In the shallow networks, with two and three three encoding layers, the NPV seems more heterogeneous compared to deeper networks. This could possibly explained by the less extended receptive field of the DL network, not covering the entire UF in this case. Therefore, the DL algorithm will not process the entire UF during the prediction, entailing more variation in intensities within smaller areas.

B.2 Kernel density estimation of pixel intensities CE-T1w scans

Figure B.2.

B.3 Agreement in qualitative evaluation of technical success

Table B.1 for absolute agreement, table B.2 for positive agreement, and table B.3 for negative agreement.

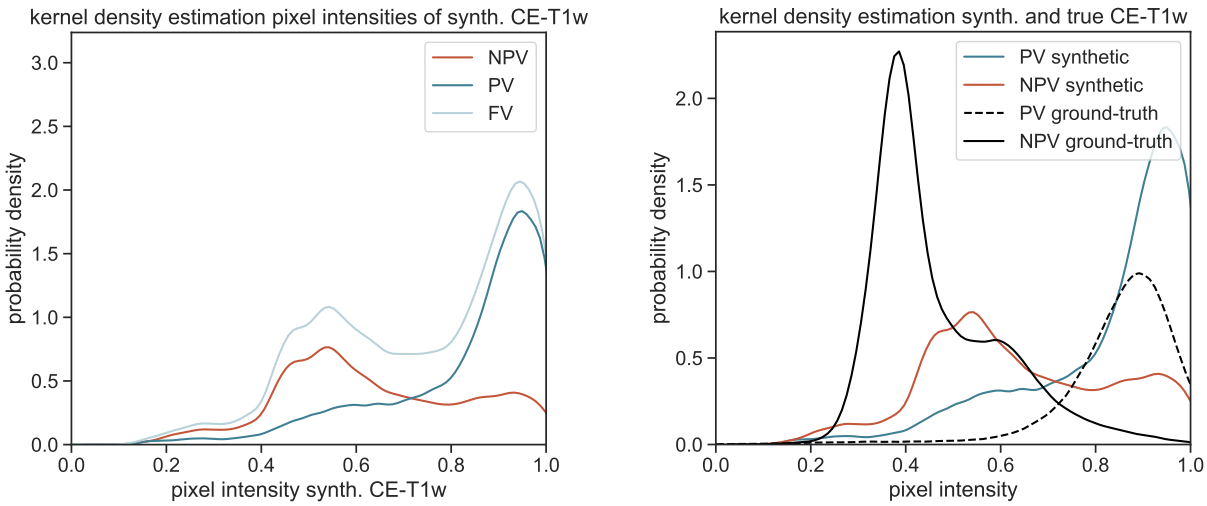


FIGURE B.2: (a) Normalized kernel density estimation of the pixel intensities of the synthetic CE-T1w scans retrieved from all the DWI datasets the test set, for the three different anatomical regions (NPV, PV and FV). In this graph, the relation is visualized between the pixel within a region (NPV, PV or FV), and the probability of that pixel to have a certain intensity (x-axis). (b) Normalized kernel density estimation of the PV and NPV for both synthetic (colored) and true CE-T1w scans (black).

B.4 More details on Friedman test statistics

Table B.4.

	Fraction	%	95% CI of %
Observer 1	14/19	73.7%	[51.9;95.5]
Observer 2	13/19	68.4%	[45.4;91.4]
Observer 3	13/19	68.4%	[45.4;91.4]
Observer 4	17/19	89.5%	[68.6;97.1]
All observers	57/76	75.0%	[64.2;83.4]

TABLE B.1: Absolute agreement between the true and synthetic CE-T1w scans, including 95%-CIs for all the observers, on all cases in the test set. These results include also the scans where no complete protocol was available, e.g. without screening scans. Please note that absolute agreement in the cases with a complete scan protocol was 83%.

	Fraction	%	95% CI of %
Observer 1	9-14	64.3%	[35.6;93.0]
Observer 2	12-18	66.7%	[42.5;90.8]
Observer 3	11-17	64.7%	[39.4;90.0]
Observer 4	13/15	86.7%	[62.1;96.3]
All observers	45/64	70.3%	[58.2;80.1]

TABLE B.2: Positive agreement between the true and synthetic CE-T1w scans, including 95%-CIs for all the observers, on all cases in the test set.

	Fraction	%	95% CI of %
Observer 1	5/10	50.0%	[12.3;87.7]
Observer 2	1/7	14.3%	[-20.7;49.2]
Observer 3	2/8	25.0%	[-13.7;63.7]
Observer 4	4/6	66.7%	[0.3;90.3]
All observers	12/31	38.7%	[23.7;56.2]

TABLE B.3: Negative agreement between the true and synthetic CE-T1w scans, including 95%-CIs for all the observers, on all cases in the test set.

		25th percentile	50th percentile (median)	75th percentile	Mean rank	Chi-square	Exact sig.
Observer 1	**Estimation NPV fraction on CE-T1w**	50.00	85.00	90.00	2.18	5,840	.052
	Estimation NPV fraction on sCE-T1w	50.00	70.00	80.00	1.55		
	Ground-truth NPV fraction	32.07	91.04	97.83	2.26		
Observer 2	**Estimation NPV fraction on CE-T1w**	65.00	80.00	90.00	2.08	1,726	.437
	Estimation NPV fraction on sCE-T1w	50.00	80.00	90.00	1.76		
	Ground-truth NPV fraction	32.07	91.04	97.83	2.16		
Observer 3	**Estimation NPV fraction on CE-T1w**	50.00	75.00	90.00	2.11	4,333	.115
	Estimation NPV fraction on sCE-T1w	50.00	70.00	80.00	1.63		
	Ground-truth NPV fraction	32.07	91.04	97.83	2.26		
Observer 4	**Estimation NPV fraction on CE-T1w**	50.00	90.00	90.00	2.00	.114	.951
	Estimation NPV fraction on sCE-T1w	40.00	80.00	90.00	1.95		
	Ground-truth NPV fraction	32.07	91.04	97.83	2.05		

TABLE B.4: More details on Friedman test including descriptive statistics, mean ranks and Friedman test outcomes. No statistical differences have been found between the estimated NPV rations based on the true and synthetic CE-T1w scans, and the ground-truth NPV ratios, as measured in 3D on the true CE-T1w scan.

Bibliography

(1) Pavone, D.; Clemenza, S.; Sorbi, F.; Fambrini, M.; Petraglia, F. Epidemiology and Risk Factors of Uterine Fibroids, 2018.

(2) Sparic, R.; Mirkovic, L.; Malvasi, A.; Tinelli, A. Epidemiology of uterine myomas: A review, 2016.

(3) Cramer, S. F.; Patel, A. *American Journal of Clinical Pathology* **1990**, *94*, 435–438.

(4) Downes, E.; Sikirica, V.; Gilabert-Estelles, J.; Bolge, S. C.; Dodd, S. L.; Maroulis, C.; Subramanian, D. *European Journal of Obstetrics and Gynecology and Reproductive Biology* **2010**, *152*, 96–102.

(5) Gupta, S.; Jose, J.; Manyonda, I. Clinical presentation of fibroids, 2008.

(6) Knight, J.; Falcone, T. Tissue extraction by morcellation: A clinical dilemma, 2014.

(7) Volkers, N. A.; Hehenkamp, W. J.; Smit, P.; Ankum, W. M.; Reekers, J. A.; Birnie, E. *Journal of Vascular and Interventional Radiology* **2008**, *19*, 1007–1016.

(8) Cardozo, E. R.; Clark, A. D.; Banks, N. K.; Henne, M. B.; Stegmann, B. J.; Segars, J. H. In *American Journal of Obstetrics and Gynecology*, Mosby Inc.: 2012; Vol. 206, pp 1–211.

(9) Borah, B. J.; Nicholson, W. K.; Bradley, L.; Stewart, E. A. *American Journal of Obstetrics and Gynecology* **2013**, *209*, 1–319.

(10) Marshall, L. M.; Spiegelman, D.; Barbieri, R. L.; Goldman, M. B.; Manson, J. E.; Colditz, G. A.; Willett, W. C.; Hunter, D. J. *Obstetrics and Gynecology* **1997**, *90*, 967–973.

(11) Parker, W. H. *Fertility and Sterility* **2007**, *87*, 725–736.

(12) Laughlin, S. K.; Schroeder, J. C.; Day Baird, D., DOI: 10.1055/s-0030-1251477.

(13) Laughlin, S. K.; Baird, D. D.; Savitz, D. A.; Herring, A. H.; Hartmann, K. E. *Obstetrics & Gynecology* **2009**, *113*, 630–635.

(14) Velez Edwards, D. R.; Baird, D. D.; Hartmann, K. E. *American Journal of Epidemiology* **2013**, *178*, 426–433.

(15) Kjerulff, K. H.; Langenberg, P.; Seidman, J. D.; Stolley, P. D.; Guzinski, G. M. *Journal of Reproductive Medicine for the Obstetrician and Gynecologist* **1996**, *41*, 483–490.

(16) Moravek, M. B.; Bulun, S. E. Endocrinology of uterine fibroids: Steroid hormones, stem cells, and genetic contribution, 2015.

(17) Cermik, D.; Arici, A.; Taylor, H. S. *Fertility and Sterility* **2002**, *78*, 979–984.

(18) Borahay, M. A.; Asoglu, M. R.; Mas, A.; Adam, S.; Kilic, G. S.; Al-Hendy, A. Estrogen Receptors and Signaling in Fibroids: Role in Pathobiology and Therapeutic Implications, 2017.

(19) Cook, J. D.; Walker, C. L. Treatment strategies for uterine leiomyoma: The role of hormonal modulation, 2004.

(20) Grantham, J. P.; Henneberg, M. *PLoS ONE* **2014**, *9*, ed. by Rosenfeld, C. S., e99776.

(21) Baron, J. A.; Vecchia, C. L.; Levi, F. *American Journal of Obstetrics and Gynecology* **1990**, *162*, 502–514.

(22) Chiaffarino, F.; Ricci, E.; Cipriani, S.; Chiantera, V.; Parazzini, F. *European Journal of Obstetrics and Gynecology and Reproductive Biology* **2016**, *197*, 63–71.

(23) Schwartz, S. M.; Marshall, L. M.; Baird, D. D. *Environmental Health Perspectives* **2000**, *108*, 821–827.

(24) Marshall, L. M.; Spiegelman, D.; Manson, J. A. E.; Goldman, M. B.; Barbieri, R. L.; Stampfer, M. J.; Willett, W. C.; Hunter, D. J. *Epidemiology* **1998**, *9*, 511–517.

(25) Wise, L. A.; Palmer, J. R.; Spiegelman, D.; Harlow, B. L.; Stewart, E. A.; Adams-Campbell, L. L.; Rosenberg, L. *Epidemiology* **2005**, *16*, 346–354.

(26) Soave, I.; Marci, R. From obesity to uterine fibroids: an intricate network, 2018.

(27) Day Baird, D.; Dunson, D. B. *Epidemiology* **2003**, *14*, 247–250.

(28) Cesen-Cummings, K.; Houston, K. D.; Copland, J. A.; Moorman, V. J.; Walker, C. L.; Davis, B. J. *Journal of the Society for Gynecologic Investigation* **2003**, *10*, 11–20.

(29) Stewart, E. A.; Laughlin-Tommaso, S. K.; Catherino, W. H.; Lalitkumar, S.; Gupta, D.; Vollenhoven, B. *Nature Reviews Disease Primers* **2016**, *2*, 1–18.

(30) Santamaria, X.; Mas, A.; Cervelló, I.; Taylor, H.; Simon, C. *Human Reproduction Update* **2018**, *24*, 673–693.

(31) Chang, H. L.; Senaratne, T. N.; Zhang, L.; Szotek, P. P.; Stewart, E.; Dombkowski, D.; Preffer, F.; Donahoe, P. K.; Teixeira, J. *Reproductive Sciences* **2010**, *17*, 158–167.

(32) Croce, S.; Chibon, F. MED12 and uterine smooth muscle oncogenesis: State of the art and perspectives, 2015.

(33) Mittal, P.; Shin, Y. H.; Yatsenko, S. A.; Castro, C. A.; Surti, U.; Rajkovic, A. *Journal of Clinical Investigation* **2015**, *125*, 3280–3284.

(34) Mäkinen, N. et al. *Science* **2011**, *334*, 252–255.

(35) Borggrefe, T.; Yue, X. Interactions between subunits of the Mediator complex with gene-specific transcription factors, 2011.

(36) Al-Hendy, A.; Laknaur, A.; Diamond, M. P.; Ismail, N.; Boyer, T. G.; Halder, S. K. *Endocrinology* **2017**, *158*, 592–603.

(37) Mosimann, C.; Hausmann, G.; Basler, K. β-Catenin hits chromatin: Regulation of Wnt target gene activation, 2009.

(38) Ono, M. et al. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 17053–17058.

(39) Luo, N.; Guan, Q.; Zheng, L.; Qu, X.; Dai, H.; Cheng, Z. *Translational Research* **2014**, *163*, 232–241.

(40) Reis, F. M.; Bloise, E.; Ortiga-Carvalho, T. M. *Best Practice and Research: Clinical Obstetrics and Gynaecology* **2016**, *34*, 13–24.

(41) Islam, M. S.; Protic, O.; Stortoni, P.; Grechi, G.; Lamanna, P.; Petraglia, F.; Castellucci, M.; Ciarmela, P. *Fertility and Sterility* **2013**, *100*, 178–193.

(42) Hawkins, L. K.; Correia, K. F.; Srouji, S. S.; Hornstein, M. D.; Missmer, S. A. *Human Reproduction* **2013**, *28*, 3000–3006.

(43) Verguts, J.; Ameye, L.; Bourne, T.; Timmerman, D. *Ultrasound in Obstetrics & Gynecology* **2013**, *42*, 713–717.

(44) Moore, K. L.; Dalley, A. F.; Agur, A. M. R., *Clinically oriented anatomy*, 8th ed.; Wolters Kluwer, Ed., 2018, p 1134.

(45) Munro, M. G.; Critchley, H. O.; Broder, M. S.; Fraser, I. S. *International Journal of Gynecology and Obstetrics* **2011**, *113*, 3–13.

(46) Wamsteker, K; Emanuel, M. H.; de Kruif, J. H. *Obstetrics and gynecology* **1993**, *82*, 736–740.

(47) Lippman, S. A.; Warner, M.; Samuels, S.; Olive, D.; Vercellini, P.; Eskenazi, B. *Fertility and Sterility* **2003**, *80*, 1488–1494.

(48) Roche, O.; Chavan, N.; Aquilina, J.; Rockall, A. *Insights into imaging* **2012**, *3*, 265–275.

(49) Lim, W. H.; Cohen, S. C.; Lamaro, V. P. *BMC surgery* **2020**, *20*, 70.

(50) Guo, X. C.; Segars, J. H. The Impact and Management of Fibroids for Fertility. An Evidence-Based Approach, 2012.

(51) Cook, H.; Ezzati, M.; Segars, J. H.; McCarthy-Keith, D. The impact of uterine leiomyomas on reproductive outcomes, 2010.

(52) Horne, A. W.; Critchley, H. O. The effect of uterine fibroids on embryo implantation, 2007.

(53) Farquhar, C. *BMJ (Clinical research ed.)* **2009**, *338*, DOI: 10.1136/bmj.b126.

(54) De Vivo, A; Mancuso, A; Giacobbe, A; Savasta, L. M.; De Dominici, R; Dugo, N; Dugo, C; Vaiarelli, A *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* **2011**, *37*, 361–365.

(55) Pritts, E. A.; Parker, W. H.; Olive, D. L. *Fertility and Sterility* **2009**, *91*, 1215–1223.

(56) Casini, M. L.; Rossi, F.; Agostini, R.; Unfer, V. *Gynecological Endocrinology* **2006**, *22*, 106–109.

(57) Cantuaria, G *Obstetrics & Gynecology* **1998**, *92*, 109–112.

(58) Rashid, S. Q.; Chou, Y. H.; Tiu, C. M. Ultrasonography of Uterine Leiomyomas, 2016.

(59) Wozniak, A.; Wozniak, S. Ultrasonography of uterine leiomyomas, 2017.

(60) Bazot, M.; Cortez, A.; Darai, E.; Rouger, J.; Chopier, J.; Antoine, J.-M.; Uzan, S. *Human Reproduction* **2001**, *16*, 2427–2433.

(61) Vitner, D.; Filmer, S.; Goldstein, I.; Khatib, N.; Weiner, Z. *European Journal of Obstetrics and Gynecology and Reproductive Biology* **2013**, *171*, 143–145.

(62) Schwärzler, P.; Concin, H.; Bösch, H.; Berlinger, A.; Wohlgenannt, K.; Collins, W. P.; Bourne, T. H. *Ultrasound in Obstetrics and Gynecology* **1998**, *11*, 337–342.

(63) Tamura-Sadamori, R.; Emoto, M.; Naganuma, Y.; Hachisuga, T.; Kawarabayashi, T. *Journal of Ultrasound in Medicine* **2007**, *26*, 941–946.

(64) Di Spiezio Sardo, A.; Calagna, G.; Santangelo, F.; Zizolfi, B.; Tanos, V.; Perino, A.; De Wilde, R. L. *BioMed Research International* **2017**, *2017*, ed. by Scioscia, M., 2518396.

(65) Farquhar, C.; Ekeroma, A.; Furness, S.; Arroll, B. A systematic review of transvaginal ultrasonography, sonohysterography and hysteroscopy for the investigation of abnormal uterine bleeding in premenopausal women, 2003.

(66) Kawamura, N.; Ichimura, T.; Ito, F.; Shibata, S.; Takahashi, K.; Tsujimura, A.; Ishiko, O.; Haba, T.; Wakasa, K.; Ogita, S. *Cancer* **2002**, *94*, 1713–1720.

(67) Spencer, J. A.; Forstner, R.; Cunha, T. M.; Kinkel, K. *European radiology* **2010**, *20*, 25–35.

(68) Spielmann, A. L.; Keogh, C.; Forster, B. B.; Martin, M. L.; Machan, L. S. *AJR. American journal of roentgenology* **2006**, *187*, 1499–1504.

(69) Dueholm, M.; Lundorf, E.; Hansen, E. S.; Ledertoug, S.; Olesen, F. *American Journal of Obstetrics and Gynecology* **2002**, *186*, 409–415.

(70) Kubik-Huch, R. A.; Weston, M.; Nougaret, S.; Leonhardt, H.; Thomassin-Naggara, I.; Horta, M.; Cunha, T. M.; Maciel, C.; Rockall, A.; Forstner, R. *European radiology* **2018**, *28*, 3125–3137.

(71) Sudderuddin, S.; Helbren, E.; Telesca, M.; Williamson, R.; Rockall, A. MRI appearances of benign uterine disease, 2014.

(72) Hricak, H.; Tscholakoff, D.; Heinrichs, L.; Fisher, M. R.; Dooms, G. C.; Reinhold, C.; Jaffe, R. B. *Radiology* **1986**, *158*, 385–391.

(73) Schwartz, L. B.; Zawin, M.; Carcangiu, M. L.; Lange, R.; McCarthy, S. *Fertility and Sterility* **1998**, *70*, 580–587.

(74) Lakhman, Y.; Veeraraghavan, H.; Chaim, J.; Feier, D.; Goldman, D. A.; Moskowitz, C. S.; Nougaret, S.; Sosa, R. E.; Vargas, H. A.; Soslow, R. A.; Abu-Rustum, N. R.; Hricak, H.; Sala, E. *European radiology* **2017**, *27*, 2903–2915.

(75) Sutton, C *Bailliere's clinical obstetrics and gynaecology* **1997**, *11*, 1–22.

(76) Singh, S. S.; Belland, L. *Current medical research and opinion* **2015**, *31*, 1–12.

(77) Farris, M.; Bastianelli, C.; Rosato, E.; Brosens, I.; Benagiano, G. Uterine fibroids: An update on current and emerging medical treatment options, 2019.

(78) Goodman, A. L. *The Journal of Clinical Endocrinology & Metabolism* **1946**, *6*, 402–408.

(79) Moroni, R. M.; Martins, W. P.; Dias, S. V.; Vieira, C. S.; Ferriani, R. A.; Nastri, C. O.; Brito, L. G. *Gynecologic and obstetric investigation* **2015**, *79*, 145–152.

(80) Hodgson, R.; Bhave Chittawar, P.; Farquhar, C. *The Cochrane Database of Systematic Reviews* **2017**, *2017*, CD012846.

(81) Roux, C; Pelissier, C; Listrat, V; Kolta, S; Simonetta, C; Guignard, M; Dougados, M; Amor, B *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* **1995**, *5*, 185–190.

(82) Matta, W. H.; Shaw, R. W.; Nye, M *British journal of obstetrics and gynaecology* **1989**, *96*, 200–206.

(83) Lethaby, A.; Vollenhoven, B.; Sowter, M. *BJOG : an international journal of obstetrics and gynaecology* **2002**, *109*, 1097–1108.

(84) Ortmann, O.; Weiss, J. M.; Diedrich, K. *Reproductive biomedicine online* **2002**, *5 Suppl 1*, 1–7.

(85) Murji, A.; Whitaker, L.; Chow, T. L.; Sobel, M. L. Selective progesterone receptor modulators (SPRMs) for uterine fibroids, 2017.

(86) Singh, S. S.; Belland, L.; Leyland, N.; von Riedemann, S.; Murji, A. The past, present, and future of selective progesterone receptor modulators in the management of uterine fibroids, 2018.

(87) Stewart, E. A. *New England Journal of Medicine* **2015**, *372*, 1646–1655.

(88) Merrill, R. M. *Medical science monitor : international medical journal of experimental and clinical research* **2008**, *14*, 24–31.

(89) Syl De La Cruz, M. D.; Buchanan, E. M. *Diagnosis and Management of Uterine Fibroids*; tech. rep. 2; 2017, pp 100–107.

(90) Sesti, F.; Cosi, V.; Calonzi, F.; Ruggeri, V.; Pietropolli, A.; Di Francesco, L.; Piccione, E. *Archives of Gynecology and Obstetrics* **2014**, *290*, 485–491.

(91) Ingelsson, E.; Lundholm, C.; Johansson, A. L. V.; Altman, D. *European heart journal* **2011**, *32*, 745–750.

(92) Stewart, E. A.; Shuster, L. T.; Rocca, W. A. *Minnesota medicine* **2012**, *95*, 36–39.

(93) Camanni, M.; Bonino, L.; Delpiano, E. M.; Ferrero, B.; Migliaretti, G.; Deltetto, F. *Journal of Minimally Invasive Gynecology* **2010**, *17*, 59–65.

(94) Pritts, E. A.; Parker, W. H.; Olive, D. L. *Fertility and sterility* **2009**, *91*, 1215–1223.

(95) Yoo, E. H.; Lee, P. I.; Huh, C. Y.; Kim, D. H.; Lee, B. S.; Lee, J. K.; Kim, D. *Journal of Minimally Invasive Gynecology* **2007**, *14*, 690–697.

(96) Gupta, J. K.; Sinha, A.; Lumsden, M.; Hickey, M. In *Cochrane Database of Systematic Reviews*, Gupta, J. K., Ed.; 5; John Wiley & Sons, Ltd: Chichester, UK, 2012.

(97) Llewellyn, O.; Patel, N. R.; Mallon, D.; Quinn, S. D.; Hamady, M. Uterine Artery Embolisation for Women with Giant Versus Non-giant Uterine Fibroids: A Systematic Review and Meta-analysis, 2020.

(98) Van Der Kooij, S. M.; Hehenkamp, W. J.; Volkers, N. A.; Birnie, E.; Ankum, W. M.; Reekers, J. A. *American Journal of Obstetrics and Gynecology* **2010**, *203*, 1–105.

(99) Karlsen, K.; Hrobjartsson, A.; Korsholm, M.; Mogensen, O.; Humaidan, P.; Ravn, P. Fertility after uterine artery embolization of fibroids: a systematic review, 2018.

(100) Verpalen, I. M.; de Boer, J. P.; Linstra, M.; Pol, R. L.; Nijholt, I. M.; Moonen, C. T.; Bartels, L. W.; Franx, A.; Boomsma, M. F.; Braat, M. N. *European Radiology* **2020**, *30*, 2473–2482.

(101) Gruetzmacher, J. *Zeitschrift für Physik* **1935**, *96*, 342–349.

(102) Lynn, J. G.; Zwemer, R. L.; Chick, A. J.; Miller, A. E. *The Journal of general physiology* **1942**, *26*, 179–93.

(103) Fry, W.; Meyers, R. *Stereotactic and Functional Neurosurgery* **1962**, *22*, 315–327.

(104) Coleman, D. J.; Lizzi, F. L.; Driller, J.; Rosado, A. L.; Burgess, S.; Torpey, J. H.; Smith, M. E.; Silverman, R. H.; Yablonski, M. E.; Chang, S.; Rondeau, M. J. *Ophthalmology* **1985**, *92*, 347–353.

(105) HICKEY, R. C.; FRY, W. J.; MEYERS, R.; FRY, F. J.; BRADBURY, J. T. *JAMA Surgery* **1961**, *83*, 620–633.

(106) Zhou, Y, *Principles and Applications of Therapeutic Ultrasound in Healthcare*; CRC Press: 2015.

(107) Quadri, S. A.; Waqas, M.; Khan, I.; Khan, M. A.; Suriya, S. S.; Farooqui, M.; Fiani, B. *Neurosurgical Focus* **2018**, *44*, E16.

(108) Lee, H.-L.; Kuo, C.-C.; Tsai, J.-T.; Chen, C.-Y.; Wu, M.-H.; Chiou, J.-F. *The Journal of Bone and Joint Surgery* **2017**, *99*, 1572–1578.

(109) Zaccagna, F.; Giulia, B.; Bazzocchi, A.; Spinnato, P.; Albisinni, U.; Napoli, A.; Catalano, C. *Journal of Therapeutic Ultrasound* **2015**, *3*, O51.

(110) Elhelf, I. S.; Albahar, H.; Shah, U.; Oto, A.; Cressman, E.; Almekkawy, M. *Diagnostic and Interventional Imaging* **2018**, *99*, 349–359.

(111) Chen, J.; Chen, W.; Zhang, L.; Li, K.; Peng, S.; He, M.; Hu, L. *Ultrasonics Sonochemistry* **2015**, *27*, 671–676.

(112) Tempany, C. M.; Stewart, E. A.; McDannold, N.; Quade, B. J.; Jolesz, F. A.; Hynynen, K. *Radiology* **2003**, *226*, 897–905.

(113) Griffiths, A.; Terhaar, G.; Rivens, I.; Giussani, D.; Lees, C. *Ultraschall in der Medizin* **2012**, *33*, DOI: 10.1055/s-0031-1299407.

(114) Funaki, K.; Fukunishi, H.; Funaki, T.; Sawada, K.; Kaji, Y.; Maruo, T. *American Journal of Obstetrics and Gynecology* **2007**, *196*, 1–184.

(115) Kim, Y.-s.; Lim, H. K.; Rhim, H. *Journal of Therapeutic Ultrasound* **2015**, *3*, O97.

(116) Verpalen, I. M.; van 't Veer-ten Kate, M.; de Boer, E.; van den Hoed, R. D.; Schutte, J. M.; Dijkstra, J. R.; Franx, A.; Bartels, L. W.; Moonen, C. T.; Boomsma, M. F. *European Radiology* **2020**, *30*, 3869–3878.

(117) Kim, Y.-s.; Keserci, B.; Partanen, A.; Rhim, H.; Lim, H. K.; Park, M. J.; Köhler, M. O. *European journal of radiology* **2012**, *81*, 3652–3659.

(118) Bohlmann, M. K.; Hoellen, F.; Hunold, P.; David, M. *Geburtshilfe und Frauenheilkunde* **2014**, *74*, 139–145.

(119) Gizzo, S.; Saccardi, C.; Patrelli, T. S.; Ancona, E.; Noventa, M.; Fagherazzi, S.; Mozzanega, B.; D'Antona, D.; Nardelli, G. B. Magnetic resonance-guided focused ultrasound myomectomy: Safety, efficacy, subsequent fertility and quality-of-life improvements, a systematic review, 2014.

(120) Verpalen, I. M.; Anneveldt, K. J.; Nijholt, I. M.; Schutte, J. M.; Dijkstra, J. R.; Franx, A.; Bartels, L. W.; Moonen, C. T. W.; Edens, M. A.; Boomsma, M. F. *European journal of radiology* **2019**, *120*, 108700.

(121) Magalhã Es Peregrino, P. F.; De, M.; Messina, L.; Dos, R.; Simõ, S.; Soares-Jú, M.; Chada Baracat, E. *Clinics* **2017**, *72*, 637–641.

(122) Rabinovici, J.; David, M.; Fukunishi, H.; Morita, Y.; Gostout, B. S.; Stewart, E. A. *Fertility and Sterility* **2010**, *93*, 199–209.

(123) Li, J. S.; Wang, Y.; Chen, J. Y.; Chen, W. Z. *Scientific Reports* **2017**, *7*, 1–8.

(124) Cain-Nielsen, A. H.; Moriarty, J. P.; Stewart, E. A.; Borah, B. J. *Journal of Comparative Effectiveness Research* **2014**, *3*, 503–514.

(125) O'Sullivan, A. K.; Thompson, D.; Chu, P.; Lee, D. W.; Stewart, E. A.; Weinstein, M. C. *International Journal of Technology Assessment in Health Care* **2009**, *25*, 14–25.

(126) Haar, G. T.; Coussios, C *International journal of hyperthermia : the official journal of European Society for Hyperthermic Oncology, North American Hyperthermia Group* **2007**, *23*, 89–104.

(127) Hussey, M, *Basic Physics and Technology of Medical Diagnostic Ultrasound*; Macmillan Education, Limited: 1985.

(128) Kim, Y. S.; Rhim, H.; Min, J. C.; Hyo, K. L.; Choi, D. High-intensity focused ultrasound therapy: An overview for radiologists, 2008.

(129) Dalecki, D. *Annual Review of Biomedical Engineering* **2004**, *6*, 229–248.

(130) Carson, P. L.; Fischella, P. R.; Oughton, T. V. *Ultrasound in Medicine and Biology* **1978**, *3*, 341–350.

(131) Paek, B. W.; Vaezy, S.; Fujimoto, V.; Bailey, M.; Albanese, C. T.; Harrison, M. R.; Farmer, D. L. In *American Journal of Obstetrics and Gynecology*, Mosby Inc.: 2003; Vol. 189, pp 702–705.

(132) Jolesz, F. A.; Hynynen, K. H., *MRI-Guided Focused Ultrasound Surgery*; Taylor & Francis: 2007.

(133) Billard, B. E.; Hynynen, K.; Roemer, R. B. *Ultrasound in Medicine and Biology* **1990**, *16*, 409–420.

(134) Peters, R. T. D.; Hinks, R. S.; Henkelman, R. M. *Magnetic Resonance in Medicine* **1998**, *40*, 454–459.

(135) Köhler, M.; Enholm, J.; Mougenot, C.; Andreae, T. *Philips Healthcare* **2012**, *NetForum Team*.

(136) Pron, G. *Ontario Health Technology Assessment Series* **2015**, *15*, 1–86.

(137) Taran, F. A.; Hesley, G. K.; Gorny, K. R.; Stewart, E. A. *Fertility and Sterility* **2010**, *94*, 331–334.

(138) Duc, N. M.; Keserci, B. Review of influential clinical factors in reducing the risk of unsuccessful MRI-guided HIFU treatment outcome of uterine fibroids, 2018.

(139) Xu, Y.; Fu, Z.; Yang, L.; Huang, Z.; Chen, W.-Z.; Wang, Z. *Journal of Ultrasound in Medicine* **2015**, *34*, 2293–2303.

(140) Voogt, M. J.; Van Stralen, M.; Ikink, M. E.; Deckers, R.; Vincken, K. L.; Bartels, L. W.; Mali, W. P. M.; Van Den Bosch, M. A. *CardioVascular and Interventional Radiology* **2012**, *35*, 1205–1210.

(141) Hijnen, N. M.; Elevelt, A.; Pikkemaat, J.; Bos, C.; Bartels, L. W.; Grüll, H. *Journal of Therapeutic Ultrasound* **2013**, *1*, 1–9.

(142) Hijnen, N. M.; Elevelt, A.; Grüll, H. *Investigative Radiology* **2013**, *48*, 517–524.

(143) Messina, C. et al. *Cancers* **2020**, *12*, 1493.

(144) Van Hecke, W.; Emsell, L.; Sunaert, S., *Diffusion Tensor Imaging: A Practical Handbook*; Springer New York: 2016, pp 1–440.

(145) Koh, D.-M.; Collins, D. J. *American Journal of Roentgenology* **2007**, *188*, 1622–1635.

(146) Stejskal, E. O.; Tanner, J. E. *The Journal of Chemical Physics* **1965**, *42*, 288–292.

(147) Le Bihan, D.; Iima, M. *PLoS Biology* **2015**, *13*, DOI: 10.1371/journal.pbio.1002203.

(148) Koh, D. M.; Collins, D. J.; Orton, M. R. Intravoxel incoherent motion in body diffusion-weighted MRI: Reality and challenges, 2011.

(149) Bihan, D. L.; Turner, R. *Magnetic Resonance in Medicine* **1992**, *27*, 171–178.

(150) Iima, M.; Le Bihan, D. Clinical intravoxel incoherent motion and diffusion MR imaging: Past, present, and future, 2016.

(151) Le Bihan, D.; Breton, E.; Lallemand, D.; Grenier, P.; Cabanis, E.; Laval-Jeantet, M. *Radiology* **1986**, *161*, 401–407.

(152) Saba, L. et al. The present and future of deep learning in radiology, 2019.

(153) O'Reilly, R. C.; Munakata, Y; Frank, M. J.; Hazy, T. E., *Computational cognitive neuroscience*; PediaPress: 2012.

(154) Krizhevsky, A.; Sutskever, I.; Hinton, G. E. In *Advances in neural information processing systems*, 2012, pp 1097–1105.

(155) Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence, 2019.

(156) Amisha; Malik, P.; Pathania, M.; Rathaur, V. *Journal of Family Medicine and Primary Care* **2019**, *8*, 2328.

(157) Litjens, G.; Kooi, T.; Bejnordi, B. E.; Setio, A. A. A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J. A.; van Ginneken, B.; Sánchez, C. I. A survey on deep learning in medical image analysis, 2017.

(158) Yi, X.; Walia, E.; Babyn, P. *Medical Image Analysis* **2019**, *58*, 101552.

(159) Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; Lungren, M. P.; Ng, A. Y. **2017**.

(160) Johnstone, E.; Wyatt, J. J.; Henry, A. M.; Short, S. C.; Sebag-Montefiore, D.; Murray, L.; Kelly, C. G.; McCallum, H. M.; Speight, R. Systematic Review of Synthetic Computed Tomography Generation Methodologies for Use in Magnetic Resonance Imaging–Only Radiation Therapy, 2018.

(161) Verpalen, I. M.; Anneveldt, K. J.; Vos, P. C.; Edens, M. A.; Heijman, E.; Nijholt, I. M.; Dijkstra, J. R.; Schutte, J. M.; Franx, A.; Bartels, L. W.; Moonen, C. T.; Boomsma, M. F. *Magnetic Resonance Materials in Physics, Biology and Medicine* **2020**, DOI: 10.1007/s10334-020-00841-9.

(162) Klein, S.; Staring, M.; Murphy, K.; Viergever, M. A.; Pluim, J. P. *IEEE Transactions on Medical Imaging* **2010**, *29*, 196–205.

(163) Le Bihan, D. *NeuroImage* **2019**, *187*, 56–67.

(164) Virtanen, P. et al. *Nature Methods* **2019**, *17*, 261–272.

(165) d'Agostino, R. B. *Biometrika* **1971**, *58*, 341–348.

(166) D'Agostino, R.; Pearson, E. S. *Biometrika* **1973**, *60*, 613–622.

(167) Demšar, J. *Statistical Comparisons of Classifiers over Multiple Data Sets*; tech. rep. Jan; 2006, pp 1–30.

(168) Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A., *Image-to-Image Translation with Conditional Adversarial Networks*, 2017, pp 5967–5976.

(169) Ronneberger, O.; Fischer, P.; Brox, T. In, ed. by Navab, N.; Hornegger, J.; Wells, W. M.; Frangi, A. F., Springer International Publishing: Cham, 2015, pp 234–241.

(170) Kingma, D. P.; Ba, J. L. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, International Conference on Learning Representations, ICLR: 2015.

(171) Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. *IEEE Transactions on Image Processing* **2004**, *13*, 600–612.

(172) Wang, Z.; Simoncelli, E.; Bovik, A. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, IEEE, pp 1398–1402.

(173) Ren, X.-L.; Zhou, X.-D.; Zhang, J.; He, G.-B.; Han, Z.-H.; Zheng, M.-J.; Li, L.; Yu, M.; Wang, L. *Journal of Ultrasound in Medicine* **2007**, *26*, 201–212.

(174) Choe, Y. S.; Lee, W. M.; Choi, J. S.; Bae, J.; Eom, J. M.; Choi, E. *Obstetrics and Gynecology Science* **2019**, *62*, 258–263.

(175) Minati, L.; Węglarz, W. P. *Concepts in Magnetic Resonance Part A* **2007**, *30A*, 278–307.

(176) Jacobs, M. A.; Herskovits, E. H.; Kim, H. S. *Radiology* **2005**, *236*, 196–203.

(177) Ikink, M. E.; Voogt, M. J.; Van Den Bosch, M. A.; Nijenhuis, R. J.; Keserci, B.; Kim, Y. S.; Vincken, K. L.; Bartels, L. W. *European Radiology* **2014**, *24*, 2118–2127.

(178) Cao, M.; Qian, L.; Zhang, X.; Suo, X.; Lu, Q.; Zhao, H.; Liu, J.; Qu, J.; Zhou, Y.; Xu, J.; Suo, S. *BioMed research international* **2017**, *2017*, 3805073.

(179) Liao, D.; Xiao, Z.; Lv, F.; Chen, J.; Qiu, L. *European Journal of Radiology* **2020**, *122*, 108670.

(180) Guo, Z.; Zhang, Q.; Li, X.; Jing, Z. *PLoS ONE* **2015**, *10*, DOI: 10.1371/journal.pone.0127964.

(181) Dijkstra, H.; Oudkerk, M.; Kappert, P.; Sijens, P. E. *Magnetic Resonance Imaging* **2017**, *38*, 47–53.

(182) Bisdas, S.; Braun, C.; Skardelly, M.; Schittenhelm, J.; Teo, T. H.; Thng, C. H.; Klose, U.; Koh, T. S. *NMR in Biomedicine* **2014**, *27*, 1184–1191.

(183) Ioannidis, G. S.; Nikiforaki, K.; Karantanas, A. *Physica Medica* **2019**, *65*, 59–66.

(184) Pang, Y.; Turkbey, B.; Bernardo, M.; Kruecker, J.; Kadoury, S.; Merino, M. J.; Wood, B. J.; Pinto, P. A.; Choyke, P. L. *Magnetic Resonance in Medicine* **2013**, *69*, 553–562.

(185) Gong, E.; Pauly, J. M.; Wintermark, M.; Zaharchuk, G. *Journal of Magnetic Resonance Imaging* **2018**, *48*, 330–340.

(186) Sun, H.; Liu, X.; Feng, X.; Liu, C.; Zhu, N.; Gjerswold-Selleck, S. J.; Wei, H.-J.; Upadhyayula, P. S.; Mela, A.; Wu, C.-C.; Canoll, P. D.; Laine, A. F.; Vaughan, J. T.; Small, S. A.; Guo, J. *Proceedings - International Symposium on Biomedical Imaging* **2020**, *2020-April*, 908–912.

(187) Kleesiek, J.; Morshuis, J. N.; Isensee, F.; Deike-Hofmann, K.; Paech, D.; Kickingereder, P.; Köthe, U.; Rother, C.; Forsting, M.; Wick, W.; Bendszus, M.; Schlemmer, H.-P.; Radbruch, A. *Investigative Radiology* **2019**, *54*, 653–660.

(188) deSouza, N. M.; Winfield, J. M.; Waterton, J. C.; Weller, A.; Papoutsaki, M. V.; Doran, S. J.; Collins, D. J.; Fournier, L.; Sullivan, D.; Chenevert, T.; Jackson, A.; Boss, M.; Trattnig, S.; Liu, Y. *European Radiology* **2018**, *28*, 1118–1131.

(189) Keenan, K. E.; Peskin, A. P.; Wilmes, L. J.; Aliu, S. O.; Jones, E. F.; Li, W.; Kornak, J.; Newitt, D. C.; Hylton, N. M. *Journal of Magnetic Resonance Imaging* **2016**, *44*, 846–855.

(190) Jafar, M. M. *World Journal of Radiology* **2016**, *8*, 21.

(191) Barbieri, S.; Gurney-Champion, O. J.; Klaassen, R.; Thoeny, H. C. *Magnetic Resonance in Medicine* **2020**, *83*, 312–321.

(192) Mongan, J.; Moy, L.; Kahn, C. E. *Radiology: Artificial Intelligence* **2020**, *2*, e200029.

(193) Lescrenier, M. *SIAM Journal on Numerical Analysis* **1991**, *28*, 476–495.