Bowel motion and implications for online adaptive radiotherapy of rectal cancer

Master Thesis in Medical Technology

by

Haldis Tuva Tolås Boge

Department of Physics and Technology

University of Bergen

June 3, 2024

Acknowledgements

First of all, I want to thank my supervisors Sara Pilskog, Johanna Austrheim Hundvin and Liv Bolstad Hysing for their guidance and support throughout this study. Thank you for always being so helpful and positive to my work. Thank you for your enthusiasm, and for always sharing your knowledge and ideas with me. I highly appreciate your support throughout this year, and I am deeply grateful for each one of you.

Thank you, Helge Egil Seime Pettersen, for providing me the Python script used in this study. Furthermore, I am grateful for getting to know everyone at Haukelandsbakken 45. I will miss your company and enjoyable conversations.

I am forever grateful for my family and friends, who have supported and motivated me throughout my five years at the university. Thank you for your invaluable encouragement. I would like to give a special thanks to my fellow students. Thank you for all the good times we have shared together, filled with amusing moments and engaging conversations.

Last but not least, I am thankful for all the opportunities this project has granted me. I am grateful for all the knowledge I have gained and the wonderful people I have met along the journey of my master thesis.

Haldis Tuva Tolås Boge
Bergen, June 2024

Abstract

Purpose: Radiotherapy (RT) of rectal cancer inevitably involves irradiation of the healthy bowel due to its close proximity. Reduction of excessive dose to the bowel is part of the radiation treatment plan optimisation. However, bowel position can vary during treatment due to bowel motion and the planned dose may not be representative for the dose delivered. Inaccurate delineation during planning increases uncertainties during RT delivery. This project aimed to test an algorithm for automatic bowel segmentation (auto-segmentation), and to study bowel motion during (intrafractional) and in between (interfractional) fractions. We also examined the impact of intrafractional bowel motion and variations in bladder filling on the dose distribution in the bowel during online adaptive radiotherapy (ART) of rectal cancer. Further, we investigated the impact of intrafractional bowel motion on risk estimates for grade ≥ 3 acute bowel toxicity.

Methods: In 15 rectal cancer patients, Cone-Beam Computed Tomography (CBCT) scans were acquired before (pre) and after (post) treatment for each of the first five fractions. The bladder, the individual small bowel and colon loops and the bowel (small bowel and colon combined) were segmented on each pre- and post-CBCT for every first (fx1) and last (fx5) fraction using a software for automatic segmentation, TotalSegmentator (TS), and by two manual delineators. Segmentation performances, intra- and interfractional bowel motion were quantified using the Dice Similarity Coefficient (DSC) and 95$^{\text{th}}$ Hausdorff Distance (HD95). The impact of inaccurate auto-segmentations and intrafractional bowel motion on the dose to the bowel were qualitatively and quantitatively investigated using dose-volume histogram (DVH) analyses for adaptive treatments. The relation between bladder volume and bowel exposure was investigated for low and high dose-volumes using statistical analyses. Dose-volume data considering intrafractional bowel motion were used in toxicity risk estimation for grade ≥ 3 acute bowel toxicity using two dose-response models.

Results: Statistically significantly poorer agreements were found between the bowel segmentations by TS and manual delineations (median DSC: 0.83, median HD95: 12.50 mm) than between the manually segmentations (median DSC: 0.90, median

HD95: 7.9 mm) in post-CBCTs of fx1. The low dose-volumes for manual bowel segmentations were statistically significantly greater than the low dose-volumes for auto-segmented bowels. The intrafractional bowel motion in fx1 and fx5 (median DSC: 0.83/0.84, median HD95: 12.59 mm/13.00 mm for fx1/fx5) was statistically significantly less than the interfractional bowel motion (median DSC: 0.72, median HD95: 20.97 mm). Intrafractional bowel motion had greater impact on low dose volumes than high dose volumes. Intrafractional increase in bladder volume was statistically significantly correlated with decrease in bowel exposure for volumes exposed to at least 20.0 Gy. The number of patients with risk estimates $\geq 5\%$ for grade ≥ 3 acute bowel toxicity varied depending on the dose-response model.

Conclusions: TS did not segment the bowel in CBCTs with adequate performance, even when post-processed. Manual segmentations were able to measure intra- and interfractional bowel motion in CBCTs. Intrafractional bowel motion and increased bladder volume reduced bowel exposure. The risk estimates for grade ≥ 3 acute bowel toxicity including intrafractional motion showed substantial variations and require further investigation with treatment specific dose-response models.

Contents

Abbreviations

AE	Adverse Event
AI	Artificial Intelligence
AJCC	American Joint Committee on Cancer
ART	Adaptive Radiotherapy
ASR	Age-Standardised Rate (ASR)
CBCT	Cone Beam Computer Tomography
CNN	Convolutional neural network
CRM	Circumferential resection margin
CRT	Chemoradiotherapy
CT	Computer Tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTV	Clinical Target Volume
DSC	Dice Similarity Coefficient
DNA	Deoxyribonucleic Acid
DRE	Digital Rectal Examination
DVH	Dose-Volume Histogram
EMVI	Extramular Vascular Invasion
EORTC	European Organisation for Research and Treatment of Cancer
EPID	Electronic Portal Imaging Device
ER	EthosRecti
ESMO	European Society for Medical Oncology
EURECCA	European Registration of Cancer Care
FOV	Field of View
Fx1	Fraction one
Fx5	Fraction five
GI	Gastrointestinal
GTV	Gross Tumor Volume
H_0	Null Hypothesis
H_a	Alternative Hypothesis

HD	Hausdorff Distance
HD95	95^{th} percentile Hausdorff Distance
HR	Homologous Recombination
HRQL	Health-Related Quality of Life
HU	Hounsfield Unit
HUH	Haukeland University Hospital
ICRU	International Commission on Radiation Units and Measurements
IGRT	Image Guided Radiotherapy
IM	Internal Margin
IMRT	Intensity Modulated Radiotherapy
IQR	Interquartile Range
ITV	Internal Target Volume
keV	Kiloelectronvolt
kV	Kilovoltage
kVCBCT	Kilovoltage CBCT
LARC	Locally Advanced Rectal Cancer
LARS	Low Anterior Resection Syndrome
LCCRT	Long-Course Chemoradiotherapy
LCRT	Long Course Radiotherapy
Linac	Linear Accelerator
LQ	Linear Quadratic
MeV	Megaelectronvolt
ML	Machine Learning
MLC	Multileaf Collimator
MRF	Mesorectal Fascia
MRI	Magnetic Resonance Imaging
MV	Megavolt
MVCBCT	Megavoltage CBCT
NCI	National Cancer Institute
OAR	Organ At Risk
PDD	Percentage depth dose
PRV	Planning organ at Risk Volume
PTV	Planning Target Volume
QA	Quality Assurance
QOL	Quality Of Life
QUANTEC	Quantitative Analysis of Normal Tissue Effects in the Clinic
Q1	First Quartile

Q3	Third Quartile
RAPIDO	Rectal cancer And Preoperative Induction therapy followed by Dedicated Operation
RT	Radiotherapy
RTOG	Radiation Therapy Oncology Group
SCRT	Short-Course Radiotherapy
sCT	Synthetic CT
SM	Setup Margin
SSD	Source-to-Surface Distance
SD	Standard Deviation
SF	Survival Fraction
TME	Total Mesorectal Excision
TNT	Total Neoadjuvant Therapy
TPS	Treatment Planning System
TS	TotalSegmentator
UICC	Union for International Cancer Control
VMAT	Volumetric-Modulated Arc Therapy
V#	Volume receiving # Gy or more
V#%	Volume receiving at least #% of prescribed dose
WW	Watch-and-Wait
3DCRT	3D Conformal Radiotherapy

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The aim of curative radiotherapy (RT) is to kill or inactivate cancer cells, while simultaneously spare healthy tissue. However, irradiation of tumours in the rectum inevitably expose the adjacent pelvic organs to radiation, such as the bladder and small bowel. The small bowel is a radiosensitive organ, and may develop acute and chronic changes in gastrointestinal (GI) physiology as a consequence of radiation exposure [1]. Small bowel side effects after modern pelvic RT are, among others, diarrhoea, an acute or late common side effect that may develop one to two weeks after RT is initiated. Late side effects such as obstruction, i.e., bowel blockage, may occur weeks to years after RT is completed [1]. Perforation, i.e., a hole in the bowel wall, is example of an other severe RT induced side effect [2]. Severe side effects occur in 5% of patients treated with pelvic RT with a total maximum dose of 45-50 Gy over five weeks, and are therefore considered as rare side effects [2]. The volume of irradiated small bowel and the use of chemotherapy are other factors affecting the toxicity risk [2].

The multi-center randomised Rectal cancer And Preoperative Induction therapy followed by Dedicated Operation (RAPIDO) study from 2021 compared a standard long-course treatment with simultaneous chemotherapy for locally advanced rectal cancer (LARC) with an experimental short course treatment followed by intensified sequential chemotherapy [3]. They found that severe diarrhoea was a common complication in both treatment groups [3]. A meta-analysis from 2019 by Holyoake et al. showed that the small bowel volume receiving at least 10 Gy (V10) can predict the risk of developing grade ≥ 3 acute small bowel toxicities such as diarrhoea [4]. The Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) had earlier found that the volume of the small bowel loops receiving at least 15 Gy should be less than 120 cc to keep the grade ≥ 3 acute small bowel toxicity rate lower than 10% [1].

Methods to reduce dose to the bowel could reduce toxicities and are explored in clinical studies using either protons or adaptive radiotherapy (ART). ART is available at Haukeland University Hospital (HUH) through the Ethos™Therapy [5]. In ART, the treatment plan can be altered during treatment prior to the delivery of each RT fraction [6]. Therefore, ART may account for bowel motion during treatment. If a bowel loop is observed close to the radiation target, ART can alter the dose distribution, consequently reducing the dose to that bowel loop. This method requires knowledge of bowel motion during an RT fraction.

The bowel will naturally move due to digestion. Small bowel loops undergo large amplitude shifts when its content changes and oscillating displacements of the wall due to peristalsis [7]. Froehlich et al. found that small bowel peristalsis waves happen 10.96 ± 2.51 times per minute, with a mean amplitude of 6.65 mm $\pm$ 1.15 mm [8]. As a result of a surgery, the bowel loops' ability to move can be reduced [9]. Mobile loops may move out of the high dose regions, and can move around resulting in different parts of the bowel receiving a smaller fraction of the dose [9]. Bowel motion also affects dose-response models for bowel loops, which have to account for bowel displacements during RT delivery in order to be accurate.

Segmenting bowel loops in computer tomography (CT) images are challenging and time consuming. The task becomes even more challenging in the reduced image quality of cone-beam computer tomography (CBCT) images used in ART. Therefore, it is unrealistic to use manual segmentation of all bowel loops during ART treatments. Automatic segmentation (auto-segmentation) using computer algorithms, however, has the potential to improve the consistency and efficiency in structure segmentation [10].

1.2 Problem Statement and Objectives

The purpose of our study was to improve the understanding of bowel motion during (intrafraction) and in between (interfraction) RT fractions and evaluate the intrafractional motion's impact for adaptive treatments of rectal cancer. Our study also investigated if auto-segmentation of bowel in CBCTs could improve segmentation consistency and reduce the workload. We have the following research questions:

(i) Can artificial intelligence (AI) algorithms automatically detect and segment the bowel in CBCT images with adequate performance?

(ii) Are manual and/or automatic segmentations able to measure and compare intra- and interfractional bowel motion in CBCT images?

(iii) Does intrafractional bowel motion affect the dose distribution in the bowel during ART of rectal cancer?

(iv) Is the probability of developing grade ≥ 3 acute bowel toxicity after ART of rectal cancer affected by intrafractional bowel motion?

Chapter 2

Theory

2.1 Rectal cancer

Cancer cells are cells that grow and spread uncontrollably. An overgrowth of cells can cause a tumour such as a polyp in the bowel. A tumour may be benign, meaning that it does not spread. However, it can develop into a malignant tumour. Malignant tumours can grow into healthy tissue and/or spread via lymph or blood vessels to other parts of the body to form secondary tumours called metastases [11]. Tumours located within the rectum with a distal extension of less than 16 cm from the anal margin are defined as rectal cancer. Rectal cancers are divided into low, middle or high cancers depending on whether the cancer is located in the distal, middle or upper rectum, respectively, according to Figure 2.1. A tumour with a distal extension of $\geq$ 16 cm is classified as colon cancer [12]. Symptoms of rectal cancer are blood in the stool or changed bowel habits. Also abdominal pain, change in appetite, weight loss and weakness may be signs of rectal cancer [13].

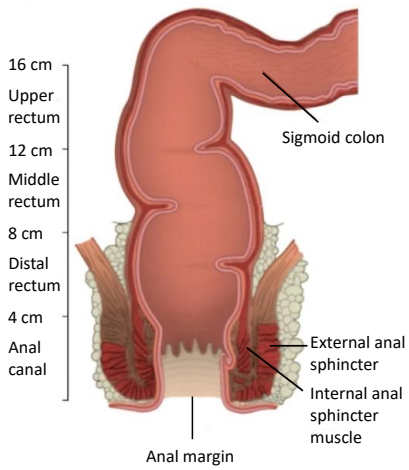

Figure 2.1: Anatomy of the rectum and anus. Illustration from [14], edited.

In 2020, the estimated number of new patients diagnosed with cancer in the rectum or rectosigmoid worldwide was 732 210 [15]. The same year, the estimated number of deaths was 339 022 [16]. This makes rectal cancer the 7th most common and the 10th most deadliest type of cancer worldwide [17]. The number of colorectal cancer incidents per year has an increasing trend, notably in developed countries. This is because of "the western way of life", associated with an unhealthy diet, obesity and lack of physical activity. Incidents of colorectal cancers are three to four times more common in developed than in developing countries [18].

In 2020, Norway had the 11th highest age-standardised rate (ASR) of rectal cancer in the world. ASR is "a summary measure of the rate that would have been observed if the population had a standard age structure" [17], calculated as a weighted mean of the age-specific rates based on the population distribution of the world's standard population. It is expressed per 100 000 person-years [17]. The ASRs of rectal cancer in Norway have been relatively stable over the past decades for both males and females. For patients over 55 years, the ASR has slightly decreased during the last 10 years. However, the ASR for patients below 55 years has slightly increased during the last 10-15 years. In 2022, 747 males and 520 females in Norway were diagnosed with cancer in the rectum. The median age at diagnosis is 71 years for males and 69 years for females. In most cases, the patients are between 60 and 80 years at diagnosis, and rectal cancer rarely occurs before the age of 40. For years 2018-2022, the five-year relative survival after treatment for patients with rectal cancer of stages (I-IV) was 73.2% [19]. In 2022, the Cancer Registry of Norway initiated a bowel screening program, with the aim of detecting colorectal cancers early to prevent cancer spread and formation of metastases. Early detected colorectal cancers are easier to cure, often resulting in fewer and less severe toxicities and a better quality of life (QOL) for the patients [20]. Figure 2.2 illustrates the relative survival of surgically treated rectal cancer patients with different cancer stages from year 2018 to 2022. Rectal cancers of stage I and II involve no spread to lymph nodes or other organs [19]. Hence, the cancer is local and less complicated to cure, entailing a better prognosis for the relative survival. A rectal cancer of stage III involves spread to lymph nodes while stage IV involves distant metastases in other organs, resulting in a worse prognosis [19].

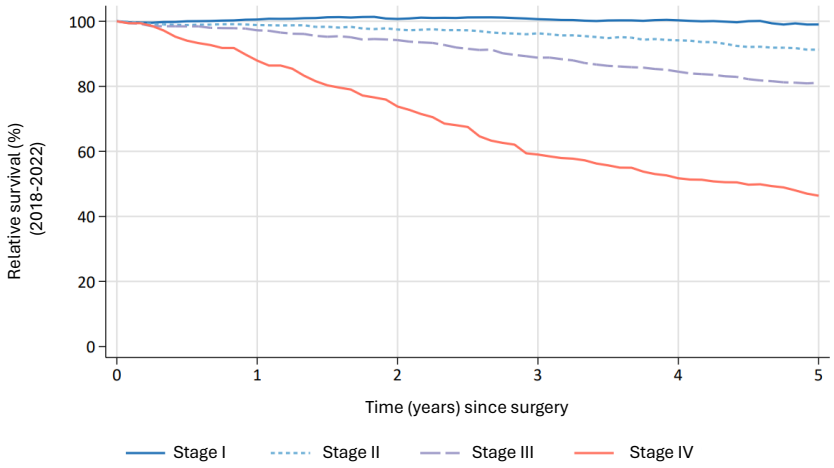

Figure 2.2: Relative survival of patients surgically treated for rectal cancers of stages I-IV in Norway from year 2018 to 2022. Illustration from [19], translated.

Classifications of rectal cancer follow the TNM staging system, from the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) to describe the size of the primary tumour and the spread of the cancer in the patient's body [21]. The primary tumour stage (T-stage) of a rectal cancer describes the depth of invasion into nearby tissue. If the tumour has extended into the submucosa, the cancer is in the T1 stage. If the tumour extends through the submucosa and into the muscularis propria, the stage is T2 [22]. A T3 staged tumour extends through the muscularis propria and into the subserosa, non-peritonealised pericolic or perirectal tissues [12]. A T4 staged tumour directly invades other organs or structures and/or perforates the visceral peritoneum [12], the peritoneum that covers the organs in the abdominal and pelvic space [23]. Figure 2.3 illustrates the different primary tumour stages of rectal cancers along with the involved surrounding anatomy.

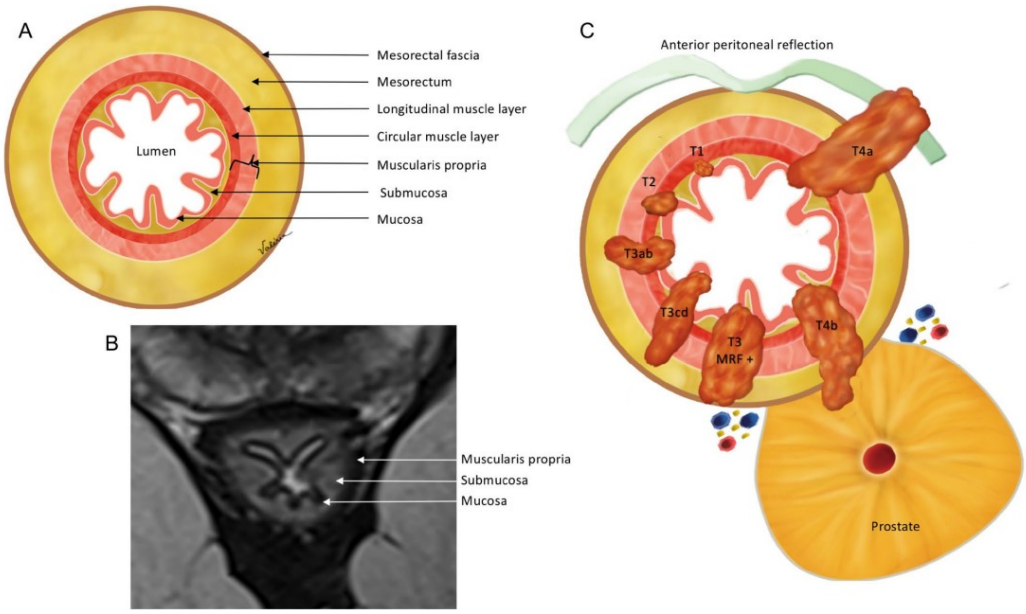

Figure 2.3: A: Illustration of the mesorectal fascia, mesorectum and rectal wall layers [24]. B: MRI of the mucosa, submucosa and muscularis propria. C: Illustration of the different primary tumour stages of rectal cancers. T3 MRF+ staged tumours invade the mesorectal fascia (MRF) or have a margin of $\leq$ 1 mm from the MRF [25]. Illustration from [24].

N-staging describes the nodal staging; the number of regional lymph nodes with metastases [22]. Lymph nodes are divided into locoregional and non-locoregional nodes. Locoregional nodes are removed in treatments involving a surgical excision. In rectal cancer, the locoregional lymph nodes are the mesorectal lymph nodes, the inferior mesenteric and superior rectal nodes, the internal iliac lymph nodes and the obturator lymph nodes. Non-locoregional lymph nodes are all other lymph nodes, including the external iliac lymph nodes, the inguinal lymph nodes, the common iliac lymph nodes and the paraaortic lymph nodes. The N-stages ranges from N0 to N2 with subgroups, depending on the number of locoregional lymph nodes with metastases. The location of the locoregional lymph nodes and some of the non-locoregional lymph nodes in the pelvis are presented in Figure 2.3 [24].

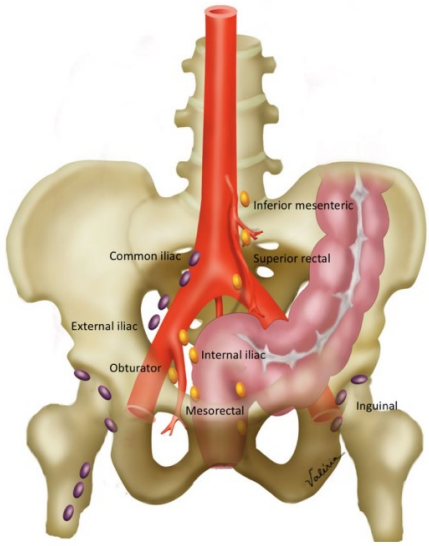

Figure 2.4: Nodal anatomy of the pelvis. Gold colored nodes illustrate locoregional lymph nodes and purple nodes illustrate non-locoregional lymph nodes for rectal cancer. Illustration from [24].

Distant metastases are described by the M-staging. Metastatic disease also refers to involved non-locoregional lymph nodes. If there are no distant metastasis, the cancer is in the stage M0, while M1 refers to the presence of distant metastases [24].

Based on the T-, N- and M-stages, rectal cancers are further grouped into different stages (0, I, II, III and IV with subgroups) [12]. The cancer stage hence describes the extent of the tumour.

In Norway, if a doctor suspects that a patient has rectal cancer, the patient is referred to a cancer patient pathway for colorectal cancer [26]. The cancer patient pathway contains national, standardised patient courses, aiming to ensure an effective initiation of further investigations and treatment [27]. The diagnosis of rectal cancer is based on findings from a digital rectal examination (DRE), which is a physical examination searching for abnormalities in the rectum [28], and endoscopy with biopsy [12]. In order to choose the right treatment, a preoperative examination of the rectal cancer is important and the cancer stage must be determined [24]. Endorectal ultrasound is utilised to assess tumours of stages T1 and T2, while magnetic resonance imaging (MRI) is more accurate in examining tumours of more complicated stages since it is able to determine the extent of the primary tumour [24, 26]. Also, CT scans of the thorax, abdomen and pelvis are utilised to gain information about the tumour's location and M-stage [26, 29]. The clinical evaluation and staging assessment may also include blood samples, among others, to investigate the presence of tumour markers [12, 29]. Findings from the clinical investigations and the cancer staging are used to assess whether the treatment has curative or palliative intent, and whether the patient

should be treated with surgery, chemotherapy, RT or a combination [30]. RT is normally given to treat rectal cancers of stage II, III and IV [26].

2.2 Rectal cancer treatment

Rectal cancer implies both small early tumours that can be directly surgically removed and larger tumours with high risk of recurrence which require preoperative treatment [26]. The main treatments for rectal cancer are primary surgery and preoperative (neoadjuvant) chemoradiotherapy (CRT) (or chemo- or RT) before surgery of the primary tumour and metastases in case of synchronous metastases [30]. Neoadjuvant treatment aims to shrink the tumour and sterilise drainage areas connected to the tumour and adjacent structures to increase the chance of successful surgery and to reduce the risk of recurrence and metastases [26, 31].

The choice between primary surgery or neoadjuvant treatment is based on various factors such as the tumour's location within the rectum and the T-, N-, and M-stages [26], as presented in Table 2.1. Tumours of stage T2 or lower combined with N0 and M0 are defined as early rectal cancers, and may be treated surgically without neoadjuvant treatment [24]. Neoadjuvant treatment is delivered as concurrent chemotherapy and long-course RT (LCCRT), which is chemotherapy delivered at the same time as long-course RT (LCRT), or short-course RT (SCRT) with or without sequential chemotherapy [26]. During surgery, the primary tumour and its lymphatic dissemination route are removed through the recommended surgical technique total mesorectal excision (TME) [26]. In TME, the mesorectum is resected and the rectum with nearby vessels and lymph nodes are removed such that the tumour and the lymph drainage for the tumour cells, which are possible routes of spread, are removed, while structures outside the rectal fascia are attempted preserved [24, 32]. The surgical treatment aims to achieve R0-resection, meaning that the resection margin has no residual macroscopic or microscopic cancer disease [26]. Radical surgery involves removal of the tumour and the tissue that may contain cancer cells. Thus, radical surgery treats the acute disease while reducing the risk of local rectal cancer recurrence [33]. When radical surgery is to be avoided, the main treatment choices are palliative stent treatment, palliative surgery, chemo- or RT, and RT may be delivered to limit the tumour's growth [29, 30]. Table 2.1 gives an overview of different treatments and techniques used to treat rectal cancer of different stages.

Table 2.1: Summary of rectal cancer stage 0-IV treatment options.

Rectal Cancer Stage (TNM classification)	Rectal Cancer Treatment
0 (Tis, N0, M0)	Surgery of a tumour or a polyp [13]. The tumour and parts of or the entire rectum, locoregional lymph nodes and fatty tissue are removed to reduce the risk of local recurrence [29].
I (T1, T2, N0, M0)	Surgery with or without chemo- and/or RT (delivered when an increased depth of tumour invasion increases the risk of spread to lymph nodes) [13]. T2 tumours in the lower rectum are often treated with preoperative RT due to the short distance to the circumferential resection margin (CRM) [26].
II (T3, T4, N0, M0) and III (Any T, N1, N2, M0)	Neoadjuvant CRT, SCRT before chemotherapy and surgery (based on the RAPIDO study) or surgery alone [26]. Internationally, the watch-and-wait (WW) approach is an option [13], but not in Norway due to the Norwait scandal [34].
IV (Any T, Any N, M1)	Treatment depends on the location and ingrowth of the primary tumour, the extent of the spread and the location of the metastasis [13]. Synchronous metastases can be treated with pre- or postoperative chemotherapy [26]. If the primary tumour is treated with RT, SCRT of 5×5 Gy is preferred as van Dijk et al. showed that radical surgical treatment for all tumour sites in patients with stage IV primary metastasised rectal cancer after SCRT and chemotherapy is a potentially curative treatment [35].

The different rectal cancer treatments have different risks of complications. Surgery may induce severe side effects such as anastomotic leak (bowel contents leaking into the body) and wound breakdown (a cut reopens), and/or more common complications such as hernia, ileus and functional disorders related to stools, urination and sexual functions, or a permanent stoma [29, 30]. Functional problems following a rectal resection are summarised under the term "low anterior resection syndrome" (LARS). LARS involves different symptoms which can reduce the patient's QOL, such as unpredictable bowel function and incontinence (inability to control the bowel movements), which results in dependency of having a toilet nearby or pad wearing, affecting both mental and emotional wellbeing. There are multiple treatments available to control LARS symptoms and over time increase the patient's QOL [36]. Chemotherapy can cause severe complications such as severe diarrhoea, and/or more common complications such as nausea, vomiting and neurotoxicity [30]. Rectal cancer patients treated with RT may experience side effects such as diarrhoea, bladder irritation, bowel inflammation, urge (sudden need to deliver stools) and incontinence (see Section 2.10) [30]. Peeters et al. found that preoperative SCRT of 5×5 Gy before TME resulted in increased local control, but also increased long-term bowel dysfunction, including among others rectal bleeding and fecal incontinence compared to patients who underwent TME alone [37]. In the RAPIDO trial, severe toxicities occurred more frequently in the patients re-

ceiving total neoadjuvant treatment (TNT) compared to the patients receiving standard LCCRT [3]. Still, the TNT tested in their trial is included in the Norwegian national guidelines for treatments of rectal cancer [26].

2.2.1 The use of radiotherapy of rectal cancer

RT of rectal cancer can be delivered pre- or postoperative as LCRT or SCRT. Preoperative therapy is preferred over postoperative therapy due to patients showing better tolerance to the treatment, less toxicity and better outcomes in terms of local control and distant metastases [26, 38, 39, 40, 41]. Preoperative RT is recommended by the Norwegian national guidelines if the tumour grows into nearby organs, if the tumorous tissue is located ≤ 2 mm from the mesorectal fascia (MRF) or ≤ 1 mm from a diseased lymph node, or if extramural vascular invasion (EMVI+) where a blood vessel is invaded or tumour deposits along draining veins have been detected [26]. About 30-40% of all rectal cancer patients in Norway have during the last years received preoperative RT [19]. The proportion of patients receiving preoperative SCRT is increasing. In 2022, 26% of all rectal cancer patients received preoperative SCRT while 11% received preoperative LCRT [19].

SCRT delivered as 5×5 Gy is given as neoadjuvant treatment as the randomised controlled TME trial by van Gijn et al. showed a reduction in 10-year local recurrence by more than 50% in patients with resectable rectal cancer without distant metastases treated with SCRT before TME compared to patients treated with TME alone [42]. LCCRT consists of RT delivered in 25 fractions with 2 Gy per fraction or in 28 fractions with 1.8 Gy per fraction combined with 5-FU based chemotherapy [26]. LCCRT showed improved local control and overall survival, among other results, in patients with nonresectable rectal cancer compared to LCRT alone [43]. The multicentre, randomised Stockholm III phase 3 trial showed no significant difference in cumulative incidence of local recurrence, distant metastases, overall survival or intercurrent death in patients treated with SCRT before surgery within 1 week, before surgery after 4-8 weeks or LCRT before surgery after 4-8 weeks (no patients received neoadjuvant LCCRT, however some patients in all three treatment groups received adjuvant chemotherapy) [44]. The European Registration of Cancer Care (EURECCA) reviewed literature on preoperative SCRT and LCCRT of rectal cancer, and found that overall survival, disease free survival and local control is approximately similar in SCRT and LCCRT of intermediate advanced rectal cancer [45].

According to the European Society for Medical Oncology (ESMO), a tumour with a T stage > cT3b, i.e., > 6-15 mm depth of invasion beyond the muscularis propria, and with an EMVI+, are considered to be locally advanced [12]. LARC patients have

high risks of developing distant metastases [26]. In Norway, approximately 10-15% of all rectal cancer patients have LARC [26]. ESMO recommends both preoperative SCRT and LCCRT as LARC treatment [46]. Earlier, the standard treatment of LARC in Norway was preoperative RT with a tumour dose $\leq 50 - 50.4$ Gy together with concomitant chemotherapy of 5FU or capecitabine [26]. In recent years, the majority of the patients in Norway receive SCRT [19], due to the results of the RAPIDO study and other similar studies [26]. RAPIDO compared standard LCCRT of LARC patients with a sequential treatment of SCRT followed by intensified chemotherapy before surgery (TNT). The study showed a reduced probability of disease-related treatment failure in the TNT group compared to the standard of care group [3]. Tolerance to TNT, however, is an issue. In the standard of care group, 7% of the patients receiving neoadjuvant treatment and 32% of the patients receiving adjuvant chemotherapy had to interrupt the chemotherapy before its completion due to toxicities [3]. In the TNT group, 14% of the patients did not complete the chemotherapy due to toxicities [3]. Within six months from completion of the neoadjuvant treatment, 89% of the patients in the standard of care group had surgery with curative intent, compared to 92% of the patients in the TNT group [3]. The 3-year overall survival was 88.8% and 89.1% in the standard and TNT group respectively [47]. The results from the RAPIDO trial suggest that the experimental treatment may be considered as a new standard treatment for patients with high-risk LARC [3]. However, the trial has received some critical comments [48]. The long-term side effects of SCRT and whether the experimental treatment from the RAPIDO trial is to be considered as the standard treatment of LARC are debated [3, 48].

ESMO mentions watch-and-wait (WW) as an alternative to surgery in LARC patients achieving complete clinical response after neoadjuvant treatment [46]. The WW approach consists of active surveillance instead of surgery, aiming to preserve the patient's organs and QOL. If the patient later experiences local regrowth, surgery must be performed [49]. In Norway, the WW approach was tested in a prospective multicenter study starting in 2018 [26]. However, the study was interrupted in November 2020 due to the Norwait scandal [34], and is therefore not used [26].

2.3 The basic physics of radiotherapy

The aim of curative radiotherapy is to use ionising radiation to kill or sterilise cancer cells while spare normal tissue and minimise complications. The two main types of RT are external beam RT and internal radiation [50].

2.3.1 External radiotherapy

External beam RT is primarily a local treatment. A machine, normally a linear accelerator (linac), aims radiation at a specific part of the patient's body where the cancer is located, and the radiation is delivered from outside the body [51]. The linacs used in RT are cyclic accelerators that normally accelerates electrons from 4 to 25 mega-electronvolt (MeV) through a linear tube using high-frequency electromagnetic waves [52, 53]. The high-energy electron beam can be used in electron therapy. For photon therapy, the accelerated electrons are directed towards a target in the linac's treatment head [52]. The target consists of a material of high atomic number when the electron beam is accelerated to energies below 15 MeV, and a low atomic number for electron energies above 15 MeV [53]. When high-speed electrons interact with a nucleus, Coulomb forces of attraction deflect the electrons from their paths, resulting in an energy loss and hence the production of bremsstrahlung x-rays. Multiple interactions create a spectrum of x-ray energies with an average photon energy of approximately one-third of the maximum electron energy [52]. Different types of linacs can accelerate electrons and provide x-rays in various energy ranges. As an example, a typical modern high-energy linac can accelerate electrons to several electron energies such as 6, 9, 12, 16 and 22 MeV, while providing photon energies of 6 and 18 megavolts (MV) [53].

2.3.2 Interactions of photons with matter

When a photon beam travels through matter, it interacts with the atoms in the matter and attenuates [52]. Hence, the photon intensity is reduced. A photon may be completely absorbed by the atom or scattered from its original path [53]. The intensity of the attenuated photon beam, $I(x)$, after traveling a distance x through the matter is given by Equation 2.1 where I_0 is the initial intensity of the photon beam and μ is the linear attenuation coefficient which depends on the energy of the photons and the atomic number and the density of the matter [52].

$$I(x) = I_0 e^{-\mu x} \tag{2.1}$$

The mass attenuation coefficient is the linear attenuation coefficient (μ) divided by density (ρ); $\frac{\mu}{\rho}$ [52], and can be interpreted as a measure of the probability for any interaction to happen between a photon and an atom in the matter, i.e., the absorbing material [53]. The probability depends on the energy of the incident photon and the atomic number for the absorbing material. In RT, the photon interactions with matter

of interest are photoelectric effect, Compton scattering and pair production [52]. At therapeutic energies, the total mass attenuation coefficient $\frac{\mu}{\rho}$ for photon interactions with matter is a combination of the probability of each interaction to happen, given by Equation 2.2.

$$\frac{\mu}{\rho} = \frac{\tau}{\rho} + \frac{\sigma_C}{\rho} + \frac{\pi}{\rho} \tag{2.2}$$

Photoelectric effect describes the process where an incoming photon is absorbed by an atom in the absorbing material, resulting in an ejection of one of the atom's bound orbital electrons [52]. The kinetic energy (E_k) of the ejected electron equals the difference between the energy of the incoming photon (hv_0) and the binding energy of the electron (E_B). When a bound orbital electron is ejected, a vacancy is created in the electron shell and the atom becomes excited. An outer electron can fill the vacancy resulting in the emission of a characteristic x-ray or an Auger electron and hence de-excitation of the atom, as illustrated in Figure 2.5 a). The emitted characteristic x-ray may further interact with the absorbing material, which in RT is the tissue [54]. The soft tissue in the body has low atomic numbers, and the K-shell's, which is the innermost shell of an atom, binding energy of soft tissue is relatively low, resulting in a low energy of the characteristic x-ray. Therefore, the energy deposited in the body by the characteristic x-ray is locally absorbed [52]. The mass attenuation coefficient of the photoelectric effect, $\frac{\tau}{\rho}$, is proportional to $\frac{Z^3}{{hv_0}^3}$ where Z is the atomic number of the absorber [53]. Thus, the probability of photoelectric effect to happen depends on the photon energy and the absorber's atomic number [52]. The probability of photoelectric effect is greatest when $hv_0 = E_B$, and decreases with increasing photon energy. The Compton effect becomes dominant when $hv_0 >> E_B$ [52].

Compton effect, illustrated in Figure 2.5 b), is the process where the photon interacts with a "free" orbital electron, i.e., an electron with $E_B << hv_0$ [53]. Part of the energy of the incident photon is transmitted to the electron which is then emitted as a Compton electron at an angle θ with respect to the direction of the incident photon. The photon is scattered at an angle ϕ. The energy of the Compton electron and the scattered photon depends on the energy of the incident photon and the scattering angle ϕ. As the Compton effect is an interaction between a photon and a free electron, it is independent of the atomic number of the absorber. The Compton mass attenuation coefficient, $\frac{\sigma_c}{\rho}$, is almost the same for all materials because it depends on the electron density of the absorbing material which is approximately the same for most materials except hydrogen [52].

When the incident photon energy is greater than 1.02 MeV, pair production illustrated in Figure 2.5 c) can happen. In pair production, the photon interacts with the

electromagnetic field of an atomic nucleus and creates an electron-positron pair with rest mass energies of 0.51 MeV, i.e., the entire photon energy is converted into mass. If $hv_0 > 1.02$ MeV, which is the threshold energy for pair production, the excess energy is shared between the electron and the positron as E_k which normally results in emission of the electron and the positron [52]. The positron continues to travel through and interact with the matter and loses energy through ionisation, excitation and bremsstrahlung before it may combine with a free electron in the annihilation process, producing two annihilation photons with 0.51 MeV each. These photons are ejected back-to-back [52]. The mass attenuation coefficient for pair production, $\frac{\pi}{\rho}$ is proportional to Z, meaning that pair production is most likely to occur in absorbing materials of high atomic numbers [53].

The photon may also interact with matter through coherent scattering, illustrated in Figure 2.5 d). This process is most probable for photon energies below 10 kiloelectronvolt (keV) and for materials with high atomic number, and thus has no practical importance in RT where photons are of higher energies and since the body is composited of tissues of lower effective atomic numbers [52].

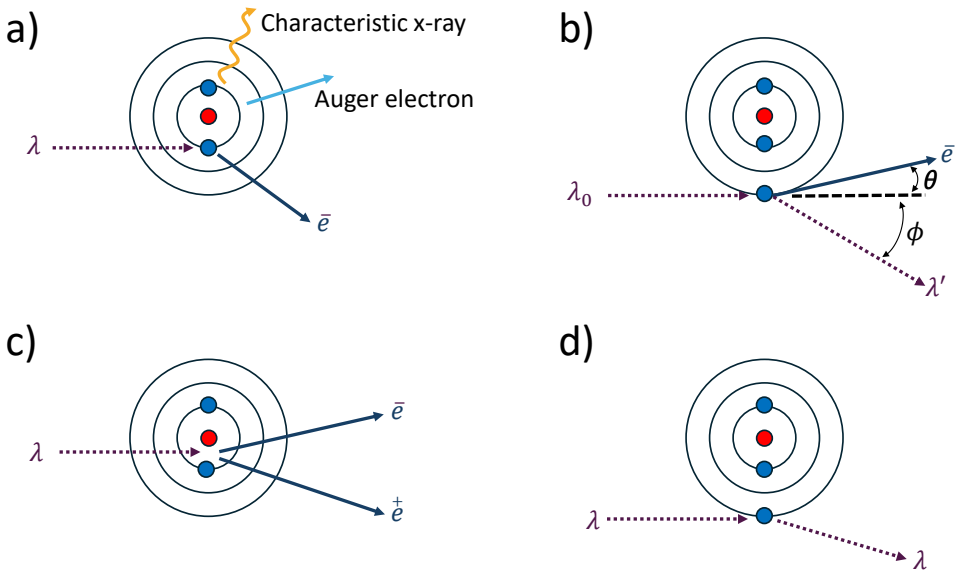

Figure 2.5: Simple illustration of photon interactions with matter. a) Photoelectric effect. b) Compton scattering. c) Pair production. d) Coherent scattering. λ: photon of energy hv. e^-: electron. e^+: positron. λ_0: incident photon. λ': scattered photon. θ and ϕ: scattering angles. Blue circles indicate electrons. Illustration inspired by [52].

The relative importance of the interactions of photons with matter are illustrated in Figure 2.6. At the left region where $hv < 1$ MeV, the photoelectric effect is the dominant interaction. As the photon energy increases, the Compton effect becomes dominant. Pair production is the dominant interaction for high energies in the region to the right [55, p. 712]. Most of the atoms in the human body consist of low atomic

numbers. The effective atomic numbers for fat and muscles are for instance 6.46 and 7.64 respectively [52]. The Compton effect is the dominant interaction for photon energies used in RT, as illustrated in Figure 2.6, and is therefore the most important interaction which RT benefits from.

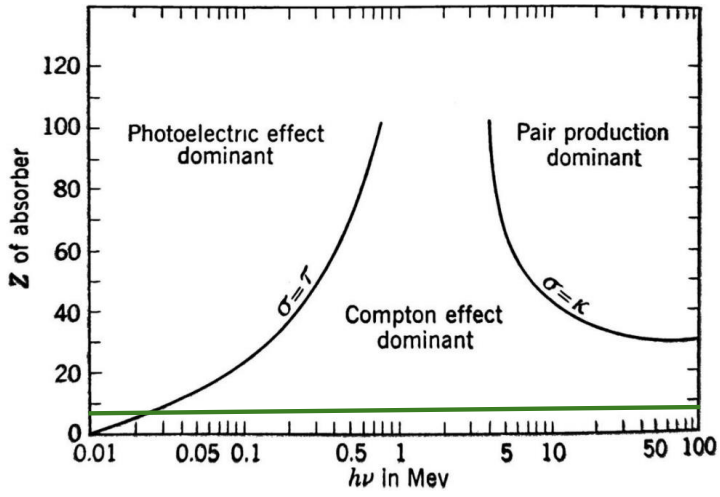

Figure 2.6: Illustration of the relative importance of photoelectric effect τ, compton effect σ and pair production κ, depending on the atomic number Z of the absorbing material and the energy hv of the photon. The relative importance of two neighboring interactions are equal at the curves [55]. The green line represents an approximate value of the effective atomic numbers for human fat and muscles. Illustration from [55], edited.

Photons are indirectly ionising radiation. When a photon beam travels through an absorber, the photons lose some/all of their initial energy, and electrons from the atoms of the absorber are ejected, depending on the interaction process as previously described. Ejected electrons travel further through the absorber, and are scattered multiple times and therefore rapidly change direction due to their low mass. Electrons interact with matter mainly by ionisation and excitation of atoms via the Coulomb force, which therefore are the most important processes resulting in energy deposition in tissue during RT [52]. A collision between an orbital electron with binding energy less than the kinetic energy of the incoming electron is an interaction between the electrons' electromagnetic fields [52]. These interactions happen in small increments along the incident electron's ionisation path. In each interaction, the electron loses some/all of its energy to the orbital electron, and thus is directly ionising. Either the orbital electron is ejected from the atom or it will be raised to higher energy levels, i.e., excitation of the atom. An ejected electron is called a δ ray if its energy is sufficient to produce a secondary ionisation path and generate more secondary electrons [52, 54]. This causes a chain reaction where multiple atoms and molecules in the absorber are ionised and excited. An electron can also interact with the electromagnetic field of a nucleus. Then, the travelling electron will experience a sudden change in direction and speed, and thus

lose parts of its energy as bremsstrahlung radiation. At very high energies, most of the electron's energy deposition happens through bremsstrahlung radiation [54].

Absorbed dose

Photons liberate electrons which deliver dose to the tissue. Absorbed dose, D, is a measure of the energy absorbed per unit mass of any absorbing material according to Equation 2.3, where $d\bar{E}$ is the mean energy transferred to a mass dm of the absorbing material. Absorbed dose has the unit gray, Gy, where $1Gy = 1J/kg$ [52].

$$D = \frac{d\bar{E}}{dm} \tag{2.3}$$

A dept dose curve explains how a radiation dose is absorbed at different depths in the absorber. Percentage depth dose (PDD) curves for photon energies of 6 and 15 MV are illustrated in Figure 2.7. When a photon beam enters the body, some photons may deflect and backscatter which, in addition to photons liberating electrons in and near the skin surface, causes entrance dose to be deposited. The amount of deposited dose increases from the entrance dose to the maximum dose (d_{max} at PDD = 100%) for increasing depth because more electrons are liberated. After the energy deposition has reached its maximum, the dose decreases for increasing depth due to attenuation of the photon beams. A photon beam of high energy results in a lower entrance dose and a higher exit dose in addition to a less steep dose reduction for increasing depths compared to a low energy photon beam. The entrance dose is decreased because deflected photons of higher energy normally scatter in a more forward direction instead of backscattering. The exit dose is increased and the dose reduction is less steep because photon beams of higher energy travelling through tissue are less attenuated compared to lower energy photons. Also, a higher photon energy yields liberated electrons of higher energy that can transfer energy at a greater tissue depth, resulting in an increased depth of d_{max}. A depth dose curve is also dependent on other factors such as the field size used during treatment delivery and the source-to-surface distance (SSD) between the gantry and the patient's surface skin [56].

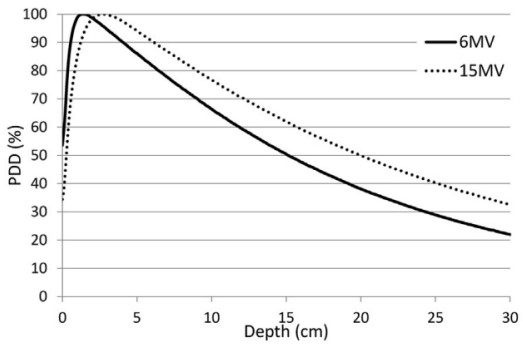

Figure 2.7: Depth dose curves for photon percentage dept doses (PDDs) for a 6 MV (solid line) and 15 MV (dashed line) photon beam. The PDDs are generated for a field size of 10 × 10 cm with a source-to-surface distance (SSD) of 100 cm [56]. Illustration from [56].

2.4 Radiobiology

2.4.1 Cell response to radiation

The anatomy of an animal cell is illustrated in Figure 2.8. The cell nucleus is the organelle in the cell that controls the cell's activity and its functions, and contains the genetic material of the cell; deoxyribonucleic acid (DNA). The DNA molecule is a double stranded, helix shaped molecule, with strands made up of nucleotides including the bases adenine (A), guanine (G), cytosine (C) and thymine (T). Hydrogen bonds connect the strands, pairing base A with T and C with G. One DNA strand is a template for reconstructing the other. The genetic information of a cell is coded according to the sequence of these bases in the DNA molecule. A cell performs cell division, called mitosis, to replace dead cells and to increase the population of cells. In mitosis, the DNA molecule is duplicated to form chromosomes. Two identical daughter cells are formed as a result of mitosis, each receiving a copy of the mother cell's DNA [57].

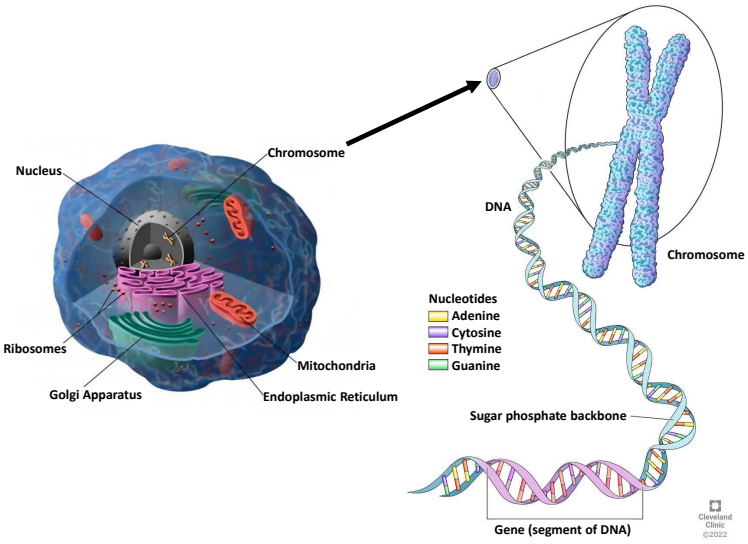

Figure 2.8: Illustration of the anatomy of an animal cell inside the cell membrane and the structure of a chromosome. Ribosomes, Golgi Apparatus, Endoplasmic Reticulum, Mitochrondria and Nucleus are organelles in the cytoplasm. A chromosome inside the nucleus consists of a protein and a DNA molecule. The nucleotides adenine, cytosine, thymine and guanine are marked in different colors. Their sequence codes for a gene, which is a segment of the DNA molecule in the chromosome [57]. Illustrations from [58] and [59], edited.

RT aims to break the chemical bonds of the DNA molecules in the cancer cells. Since a cell needs the information contained in the DNA to reproduce itself, the DNA is considered a critical molecule causing cell death if destroyed. If the DNA is exposed to radiation, it might induce different levels of damage [54]. Severe damages may cause the cell to die, while less severe damages can be repaired. If the damage is irreparable, the cell can die within hours to days. Early effects of radiation occurs when damaged cells die in the attempt to divide, resulting in early tissue reactions. Late effects of radiation occur when the damages are expressed long after the radiation, resulting in late tissue reactions such as cancer or life span shortening. Examples of radiation induced cell damages are irreparable damages such as apoptosis (programmed cell death, the cell dies before it divides), and repairable damages such as adaptive responses (the irradiated cell becomes more resistant to later radiation) and mutations (the surviving cell contains a change in the DNA sequence), among others [53]. As a photon beam penetrates the body, it leaves behind a trail of ionised and excited molecules and atoms. Only a small part of the photon energy is transferred directly to the biological sensitive material [54].

Indirect damage due to radiation

Mostly, the primary interactions between the photons and the tissue ionises simpler molecules, producing free radicals. A free radical is an electrically neutral atom or

molecule carrying an unpaired electron in the outer shell, making it chemically extreme reactive. The free radicals can reach and induce chemical changes at critical sites in biological structures [54]. A cell consists of approximately 80% water (H_2O). The water in a cell can be ionised due to photon radiation, resulting in high energy electrons being ejected from the water molecules. The ionisation of water is expressed in Equation 2.4, where the water molecule releases an electron, e^-, and creates a positively charged ion, H_2O^+. Another water molecule will most likely capture the released electron, resulting in the production of a negative ion, H_2O^-, according to Equation 2.5. The H_2O^+ and H_2O^- ions are unstable and create the free radicals $OH^\bullet$ and $H^\bullet$ as they decay according to Equation 2.6 and 2.7 respectively [54]. The dots ($\bullet$) in Equation 2.6 and 2.7 indicate that the hydroxyl (OH) and hydrogen (H) are free radicals. The free radicals can damage for instance the DNA by breaking chemical bonds, inducing cell death or genetic mutations. The majority of the DNA damages produced by ionising radiation are due to this indirect chemical damage. Mostly, when photons induce such chemical damages, the indirect damages are not permanent [54].

$$H_2O + radiation \longrightarrow H_2O^+ + e^- \tag{2.4}$$

$$e^- + H_2O \longrightarrow H_2O^- \tag{2.5}$$

$$H_2O^+ \longrightarrow H^+ + OH^\bullet \tag{2.6}$$

$$H_2O^- \longrightarrow H^\bullet + OH^- \tag{2.7}$$

The cell cycle

The cell cycle is the cell's proliferation cycle. The cell cycle illustrated in Figure 2.9 describes the phases that a cell goes through between two mitoses. The cell cycle time, which is the time between two successive mitoses, is typically longer in normal tissue cells compared to cancer cells. During tissue regeneration after injury, normal cells can proliferate faster, shortening the cell cycle time [53]. The interphase, i.e. the phase between the end of the previous mitosis and the earliest signs of the next mitosis, consists of the phases G_1, S and G_2. G_1 and G_2 are referred to as gap periods or resting phases, where the cells continue to perform their specialised functions and metabolic processes [57]. G_1 is the first phase of the cycle. Here, the DNA has not yet been synthesised. The G_1 phase is followed by the S phase, where DNA synthesis and duplication occurs. After the S phase, the cell enters the next gap phase; G_2 [53].

The *M* phase is the phase for mitotic cell division which is the last phase of a cell cycle [57]. The degree of a cell's radiosensitivity depends on which phase of the cell cycle it undergoes. Cells are most radiosensitive in the *M* and G_2 phases, assumed to be due to the cells passing a checkpoint in G_2 that blocks or slows the irradiated cells' progression into the next phase. In the *S* phase, especially in the late *S* phase, cells are most radioresistant, which is assumed to be due to homologous recombination (HR), a repair mechanism for double-strand DNA breaks. HR utilises a homologous undamaged DNA with a base sequence equal to the sequence of the damaged DNA as a template to repair the damage. In the *S* phase, the availability of the undamaged DNA template is increased, which facilitates HR. Also, during DNA synthesis and duplication, the DNA undergoes conformational changes which can facilitate easier access for repair mechanisms [9, 60]. During mitosis, the DNA molecule splits into two halves consisting of one of the two original DNA strands. One half can be used as a template to construct the other matching half, hence two DNA molecules identical to the original molecule can be created [54]. If radiation induces damage to one of the two DNA strands, i.e., a single-strand break, the DNA can be repaired using the other strand as a template. However, multiple single-strand breaks may be difficult to repair, especially if they are close and on both strands. If both strands are damaged on identical regions across from each other, i.e., a double-strand break, this may induce cell death. There are multiple repair mechanisms to repair both single- and double strand breaks. However, if the repairing induces sufficient changes to the repaired DNA compared to the original DNA, complications such as mutations can occur. Radiation induced cell death is correlated with the number of double-strand breaks, and a sufficient number of double-strand breaks must be induced to kill a cell [60].

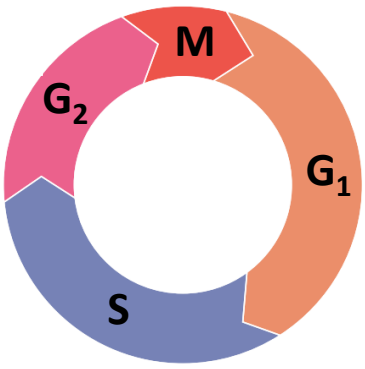

Figure 2.9: Illustration of the cell cycle with its phases. The cycle proceeds clockwise, as indicated by the arrows, starting from the first gap period G_1, through the S phase for DNA synthesis and duplication and the second gap period G_2, ending with the M phase for mitosis [53, p. 487]. Illustration from [60, p. 283].

2.4.2 Cell survival curve and the LQ-model

A cell survival curve plots the survival fraction (SF) of irradiated cells against radiation dose, illustrated in Figure 2.10 [9]. The SF is the ratio between survived irradiated cells and not irradiated control cells. Cells that maintain their reproductive capacity after radiation are surviving cells.

The shape of the cell survival curve can be different for different types of radiation, described by the Linear Quadratic (LQ) model [53]. For instance, radiation by protons result in steeper cell survival curves than photon radiation. The LQ-model gives the relation between dose and cell survival, explained by Equation 2.8, where SF is the survival fraction of cells after irradiation of dose D [53]. The LQ-model assumes that cell death due to radiation is determined by constants α and β. α describes the initial slope of the cell survival curve and the linear contribution to damage, while β describes the quadratic component and the quadratic contribution [9, 53]. α represents a single-hit kill, causing unrepairable damage to the DNA, and is independent of fractionation and dose rate. β describes a two-hit kill, causing repairable damage to the DNA, depending on the fractionation scheme used for the RT delivery and the LET of the radiation, which describes the biological effectiveness of the radiation [60]. The ratio α/β decides how bending the cell survival curve is. With reduced dose rate, the cell survival curve loses its bending shape and follows the initial slope in an approximately straight line (the dashed line in Figure 2.10), and the quadratic component of cell killing (β) approaches zero due to the single-hit kills occurring less frequently and because repair mechanisms are granted time to repair the cell. α/β gives the dose where the linear and quadratic contributions to damage and cell death are equal. In the example of Figure 2.10, $\alpha/\beta = 8$ Gy [53].

$$SF = e^{-(\alpha D+\beta D^2)} \tag{2.8}$$

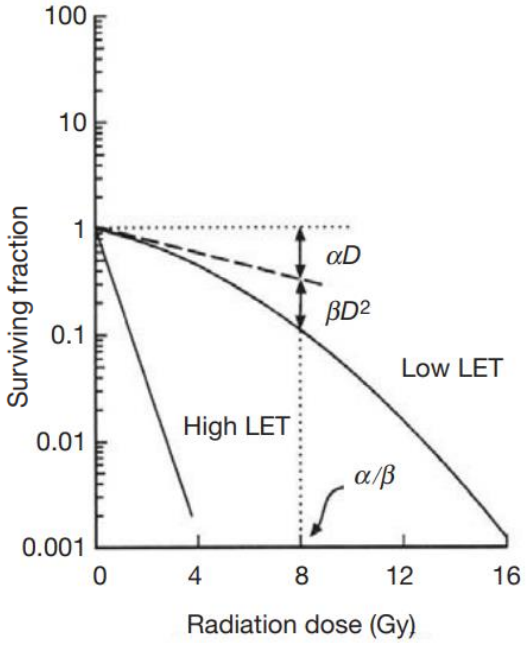

Figure 2.10: Cell survival curves for high and low LET radiation, described by the Linear Quadratic (LQ) model. Photon radiation is an example of low LET radiation. Illustration from [53].

The α/β ratio varies among different tissues, tumour types and radiation reactions, and depends, among other factors, on the tumour site and histology. In certain tissues, it may take months to years before the effects of radiation become apparent, known as late responding tissues, which have low α/β values of approximately 0.6-6 Gy. The opposite applies to tissues with high α/β values of approximately 7-20 Gy, i.e., early or acute responding tissues, where the effect of radiation becomes apparent within days to weeks after irradiation [9]. Human tumours seem to have a high α/β ratio, hence having a similar radiation response as early responding tissue [61]. The generic α/β value used for early responding tissues and tumours are 10 Gy, while late responding tissues have a generic α/β value of 3 Gy [53]. Several studies have investigated the α/β value for the bowel, resulting in a range of various values depending on the studied endpoint [9, p. 107,122]. Different α/β values are used depending on the investigated effect, i.e., a late or acute side effect. An example of cell survival curves for early and late responding tissues are illustrated in Figure 2.11. For low doses, an increase in dose does not lead to a rapid decrease in the surviving fraction for both early and late responding tissues, because DNA damages are repaired [54]. As the dose increases, the DNA damages are more severe, inducing more cell deaths. The curvature of the cell survival curves for early and late responding tissues are different due to the difference in the α/β ratio. For early responding tissues, the α/β ratio is high and α dominates at low doses, causing less bending curve in Figure 2.11. For late responding tissues, β becomes relevant at doses lower than for early responding tissues [53].

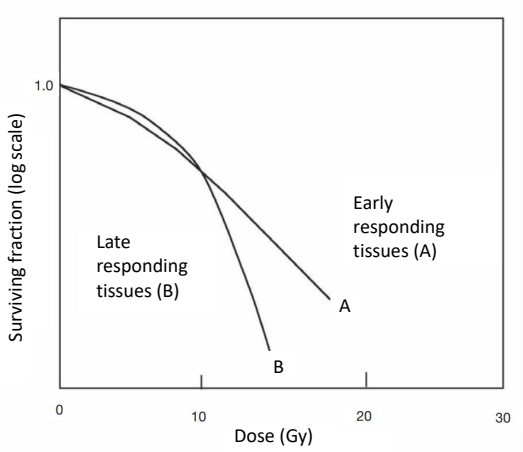

Figure 2.11: Example of cell survival curves in late (curve B) and early (curve A) responding tissues. Illustration from [53], edited.

2.5 Fractionation

Fractionation is splitting the total radiation dose into smaller doses delivered at separate time intervals. Photon therapy is more effective and induces less damage to normal tissues when the dose is delivered in multiple smaller doses instead of a single large dose [54]. Two cell survival curves are illustrated in Figure 2.12, one where the entire dose is delivered in one single fraction (A) and one where the total dose is given as a series of equal fractions (B), i.e., the dose is fractionated. For a given dose, a greater amount of the cells survive when the total dose is given in multiple smaller doses separated by time compared to as a single large dose. As Figure 2.12 illustrates, the shoulder of the cell survival curve in Figure 2.10 is repeated at each fraction (D_{1-4}) because the dose is applied at a slow rate which allows for sublethal damages to be repaired in between fractions, hence increasing the survival fraction of the normal cells [54].

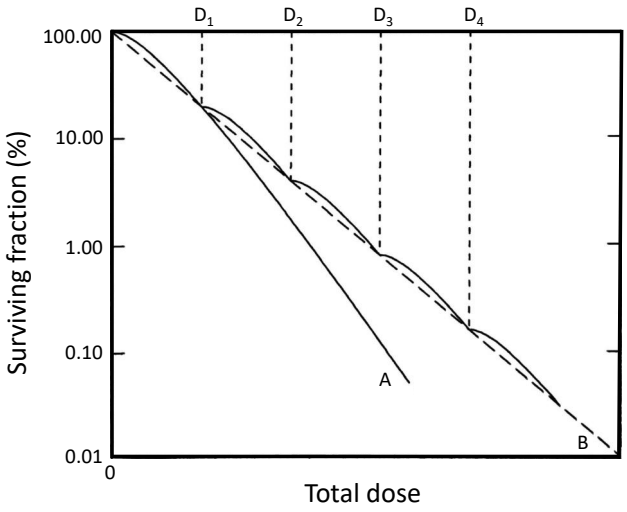

Figure 2.12: Illustration of cell survival curves. Curve A represents the cell survival curve when the total dose is delivered in one single fraction. Curve B represents the cell survival curve when the total dose, equal to the total dose in A, is delivered in four fractions (D_1, D_2, D_3, D_4), the dashed vertical lines. Illustration from [54], edited.

Fractionation has different influence on early and late responding tissues. The ratio of isoeffective doses related to different fractionation doses for early and late responding tissues are illustrated in Figure 2.13 [9]. The isoeffective doses refers to the dose needed to produce the same biologic effect when comparing for instance different fractionation schemes. The ratio of total isoeffective doses is used to compare the biological effectiveness of different radiation treatments, defined by the ratio between the doses from different treatments that produce an equal biologic effect. Conventional fractionation is the current standard fractionation scheme in RT and constitutes five daily treatments per week (ideally Monday to Friday) of 2 Gy per fraction over several weeks [9]. LCRT of rectal cancer is an example of conventional fractionation. The total dose ranges from approximately 40 Gy to 70 Gy, depending on the cancer [9]. In Figure 2.13, the ratio of total isoeffective doses are shown relative to conventional fractionation. As the figure illustrates, the ratio of total isoeffective doses decreases more rapidly for late responding tissues compared to early responding tissues when the fractionation dose is increased above 2 Gy. Decreasing the fractionation dose below 2 Gy, known as hyperfractionation, results in a higher ratio of total isoeffective doses for late responding tissues compared to early responding tissues. Therefore, changing the dose per fraction has a grater influence on late than early responding tissues [9]. Figure 2.13 illustrates that an increase in the dose per fraction above 2 Gy, known as hypofractionation, is associated with increased damage to the late responding tissue, while the damage to the early responding tissue and tumour remains relatively stable, with only a small increase. The SCRT of 5×5 Gy tested in the RAPIDO trial [3] is an example

of a hypo-fractionated RT treatment. On the other hand, with hyperfractionation, less damage is induced to both the late and early responding tissue, with greatest effect for the late responding tissue. However, reducing the fractionation dose also reduces the radiation induced damage to the tumour. Therefore, with reduced fractionation dose, the total dose must be increased in order to achieve the same tumour control as with 2 Gy fractionation. To reduce the overall treatment time, a patient can receive multiple fractions per day, known as accelerated fractionation, increasing the received average dose per week. The aim of accelerated fractionation is to reduce the repopulation of clonogenic cells, which by definition are cells that can produce an expanding family of descendants as a result of mitosis. However, accelerated fractionation increases the probability of early side effects to occur [9].

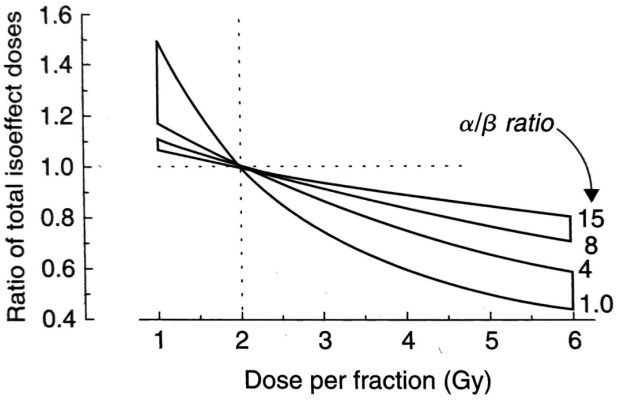

Figure 2.13: Theoretical isoeffective (isoeffect) curves based on the Linear Quadratic (LQ) model relative to conventional fractionation [9]. The range of α/β ratios between 1.0-4 and 8-15 refers to late and early responding tissues, respectively. Illustration from [9].

Equivalent total dose in 2 Gy fractions

The recommended method to compare the biological effectiveness of different treatment schedules, consisting of different total and fractionation doses, is calculating the equivalent dose in 2 Gy fractions, EQD_2. Each treatment schedule is converted into an equivalent schedule with a dose per fraction of 2 Gy. The biological effect of the equivalent schedule equals the effect of the initial schedule. The total dose of the schedule delivering 2 Gy per fraction, EQD_2, is calculated using Equation 2.9, where D is the total dose delivered through a fraction size of d Gy delivered in n numbers of fractions, and $D = n \times d$ [9]. Hence, EQD_2 is the total dose needed to induce the same biological effect as the dose D, when the dose per fraction is 2 Gy instead of d Gy.

$$EQD_2 = D\frac{d+(\alpha/\beta)}{2+(\alpha/\beta)} \tag{2.9}$$

To induce the same biological tumour effect as 50.4 Gy delivered in 1.8 Gy fractions, a total dose of approximately 40 Gy delivered in 5 Gy fractions is required, determined using the EQD_2 formula with the generic α/β value of 10 Gy for the tumour. This illustrates how hypofractionation requires less total dose to induce the same biological effect as a hyperfractionated treatment regime. The same principle applies for acute side effects, where an α/β of 10 Gy is commonly used for acute bowel toxicities.

2.5.1 The 5 R's of fractionated radiotherapy

Fractionation takes advantage of the difference in radiation response between tumour cells and normal cells, related to the 5 R's of radiotherapy [62, 63]. The aim is to achieve an optimal radiation treatment where the cancer cells are killed and the normal tissue is spared. By fractionation the treatment, this balance between the tumour and normal tissue radiation response can be achieved.

Repair refers to the cells' ability to repair radiation induced damages [53]. When the total dose is delivered through several fractions of a lower dose per fraction with sufficient time between subsequent fractions, the cells can repair damages from the previous fraction, making them more radioresistant to the next fraction. Figure 2.12 illustrates that fractionation increases the cells' surviving fraction due to factors such as repair.

Regeneration, or Repopulation, refers to the cells' ability to repopulate during the time between two subsequent fractions [53]. The dose delivered in each fraction causes a depopulation of clonogentic cells. Each fraction reduces the population of both tumour and normal clonogenic cells. However, in between fractions, survived clonogenic cells can repopulate, resulting in tumour growth. Repopulation of clonogenic tumour cells reduces the efficacy of the treatment. Therefore, in fractionation, a higher total dose is required to achieve local tumour control. Overall, the effect of all the fractions decreases the number of clonogenic tumour cells, despite the repopulation [9]. As mentioned earlier, damaged cells have a shorter cell cycle time than tumours. Hence, damaged normal tissue may repopulate faster than tumours.

Reoxygenation refers to increasing the oxygen level of the hypoxic cells, resulting in more radiosensitive cells [53]. The biological effect of radiation is affected by the amount of oxygen in the irradiated cells. Oxygen rich cells, known as aerated cells, are more radiosensitive than cells with low oxygen levels, known as hypoxic cells [54]. Tumour cells are rapidly multiplying, outgrowing their blood supply, and become hypoxic. A tumour can consist of a hypoxic core surrounded by aerated, radiosensitive cells. Each RT fraction causes some aerated cells to die. The blood supply then reaches more of the hypoxic cells, making them aerated due to reoxygenation. The newly aer-

ated cells are radiosensitive during the next fraction, and may be killed by subsequent radiation, resulting in further reoxygenation of other hypoxic cells. Overall, the tumour size will shrink, and eventually the tumour is destroyed [54].

Redistribution is related to how the cells are distributed among the various phases of the cell cycle [53]. As mentioned earlier, cells in the S phase are most radioresistant while cells in the G_2 and M phases are most radiosensitive. After radiation, cells may lose their ability to reproduce and die. This especially applies to the cells in the G_2 and M phases, while cells in the S phase mostly retain this ability. Therefore, immediately after RT delivery, most remaining cells are in the S phase. At some time after irradiation, the surviving cells are again distributed among the phases of the cell cycle [9]. When the cells are exposed to further radiation, cells that previously were in the radioresistant S phase may be in a radiosensitive phase. In this way, fractionation causes an increased cell death because the radiation are distributed over time, reaching a greater amount of cells in their radiosensitive phases.

Radiosensitivity refers to the fact that different types of tumour and normal cells have different radiosensitivities, hence responding differently to radiation [53].

2.6 Basic principles of the radiotherapy treatment

In the planning and the subsequent delivery of RT, the therapeutic goal is assured by constructing a treatment plan that fulfill a set of dose constraints for the tumour and normal tissue. RT delivery should, if feasible, assure that the treatment plan is delivered with high precision.

2.6.1 Radiotherapy treatment planning

The patient's 3D anatomy is studied to find the optimal treatment plan. CT is used to construct this 3D model of the patient and to gain information about the internal anatomy of the patient. Figure 2.14 illustrates a 3D representation of a rectal cancer patient with segmented internal structures and body contour from the patient's planning CT. Ideally, the imaging is carried out with the patient obtained in the same position as during the RT delivery. The CTs are then used to identify, localise and segment target structures, organs at risk (OARs), and other anatomical structures of interest prior to treatment [52]. They are also used to examine the dose distribution within the patient as part of the RT planning.

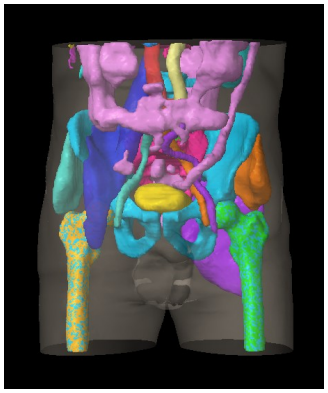

Figure 2.14: 3D representation of a rectal cancer patient based on the planning CT. The body is contoured and internal structures are segmented using the auto-segmentation algorithm TotalSegmentator.

Different organs and tissues attenuate photons to various degrees. The distributions of the linear attenuation coefficients within a CT scan are used to reconstruct an image illustrating all internal body structures. A CT image is a mathematical reconstruction based on information of how the photons are attenuated within the scanned body, performed by a reconstruction algorithm [52]. The distributions of the linear attenuation coefficients are determined by performing an inverse transformation of projection values that describes the attenuations of the photon intensities. Modern CT scanners use convolution-backprojection, a type of iterative reconstruction, to generate CT value profiles for continuously measured intensity attenuations. In addition to reconstruction, the convolution can influence the quality of the CT image, reducing image noise or increasing spatial resolution depending on the chosen convolution kernel [64]. Hence, when the distribution of the linear attenuation coefficients are determined, the CT image contains information about the density of the imaged tissue. During RT, this density affects the attenuation of the radiation beam. Therefore, CTs can be utilised in dosimetric calculations [65].

MRI is an other imaging modality used to acquire 3D volumetric images. MRI provides high-contrast images of soft tissues and are often used in combination with CTs in RT planning [52]. Since MRI images do not contain information about the density of the imaged tissue, it can not be utilised in dose calculations.

Modulated techniques are the standard, most frequently utilised treatment techniques in modern RT. In Intensity-Modulated Radiation Therapy (IMRT), the intensity of the radiation beam varies across the irradiated volume within the patient. That is, a nonuniform fluence is delivered to the patient. Modifying the intensity over multiple beam angles optimises the overall dose distribution within the patient, such that the planning target volume (PTV) is irradiated with a beam of high intensity, while the normal tissue is exposed to a beam of low intensity or no beam at all [52]. IMRT is mostly delivered using the computer-controlled multileaf collimator (MLC) technique

either by the "step-and-shoot" method or by continuously moving MLCs. MLCs consist of multiple collimating blocks or leaves made of photon absorbing materials used to shape the radiation fields [52]. The leaves can be positioned independently of each other, and are automatically controlled and positioned to generate any desired field shape. "Step-and-shoot" refers to the method where each field is divided into subfields by positioning the MLCs in different ways. The radiations from each subfields are delivered one at a time. When the MLC changes position to form the next subfield, the linac does not deliver radiation. The total beam delivered from all subfields belonging to the "main" field creates the intensity modulated beam. In IMRT delivered with continuously moving MLCs, known as dynamic MLC delivery, the linac delivers radiation all the time, even when the MLCs are moving. Thus, a variable beam intensity is delivered from different regions in the field. Wedges, which are absorbing filters used to decrease the radiation intensity, are used to shape the dose distribution of the beam [52]. In Volumetric-Modulated Arc Therapy (VMAT), a rotational cone beam of varying shape and intensity is delivered. The gantry of the linac is continuously rotating around the patient's body while the MLCs vary the intensity of the beam [52]. 3D Conformal Radiotherapy (3DCRT) is a non-modulated treatment technique where the radiation beams are of uniform intensity across the fields [52]. An IMRT plan using 12 fields with MLCs modifying the intensity profiles of the beams is illustrated in Figure E6.6. Figure E6.6 c) shows how the MLCs modify the beams from different fields to conform the fields to the PTV from various angles.

IMRT and VMAT utilise inverse treatment planning, where the clinician specifies treatment criteria such as a maximum tolerated dose for a critical structure, and an inverse planning algorithm in the treatment planning system (TPS) calculates the optimal fluence profiles and beam directions. Parameters affecting the delivered radiation intensities are adjusted to satisfy the predefined criteria. 3DCRT utilises forward planning, where the optimal design choices for the fields, beam arrangements, weights and MLCs are made on a trial-and-error basis. 3D dose distributions and dose-volume statistics calculated by the TPS are studied to determine the optimal design choices. Both forward and inverse planning require 3D image data of the patient's anatomy with delineated internal structures. Plan evaluation of both the modulated techniques and 3DCRT relies on studying dose distributions [52].

2.6.2 Evaluating the dose distribution

When constructing and evaluating an RT treatment plan, the dose distribution within the target volumes and critical organs are investigated. Therefore, the structures of interest must be defined and segmented in advance. Dose distributions can be studied

and evaluated through isodose curves in individual slices or planes from the CTs or in dose-volume histograms (DVHs). The International Commission on Radiation Units and Measurements (ICRU) have recommended a dose specification system for external photon beams. According to the ICRU Report 50, the absorbed dose of the PTV should be between 95% and 107% of the prescribed dose. According to the ICRU Report 83, the dose received by 98% of the PTV volume, ($D_{98\%}$) should equal at least 95% of the prescribed dose, and only 2% of the PTV ($D_{2\%}$) should receive more than 107% of the prescribed dose [66]. The dose to the normal tissue and OARs should be minimised. Visualising isodose curves in the CTs illustrate the anatomic locations of regions receiving uniform dose, high dose and low dose [52]. In Figure 2.15 a), isodose curves in Gy from an IMRT plan are shown within the pelvis through the sagittal plane in a rectal cancer patient. Illustrated regions receiving the highest doses, indicated by green lines in the fiure, are located within or close to the target volume, while other pelvic regions receive lower doses indicated by the outermost blue line. DVHs describe how the dose is distributed in a target volume or an OAR. DVHs can be both differential and cumulative, the latter being most important in RT and is commonly used in treatment planning and evaluating, and to compare treatment plans. A cumulative DVH illustrates the volume of a specific structure that receives a certain dose or higher [52]. DVHs can be used to extract dose-volume parameters, relating absorbed dose to volumes exposed to radiation. For instance, the volume metric V5 is the volume of a structure receiving at least 5 Gy. Figure 2.15 b) illustrates DVHs for the PTV and the bowel, which is an OAR in RT of rectal cancer, for the patient imaged in Figure 2.15 a). The PTV's DVH, indicated by the red curve, illustrates that the majority of the PTV's volume receives prescribed dose which is 25 Gy. The bowel's DVH, indicated by the curve in magenta, shows that a great amount of the bowel receives low doses (approximately < 10 Gy), and the volume of the bowel exposed to high dose is decreasing for increasing dose.

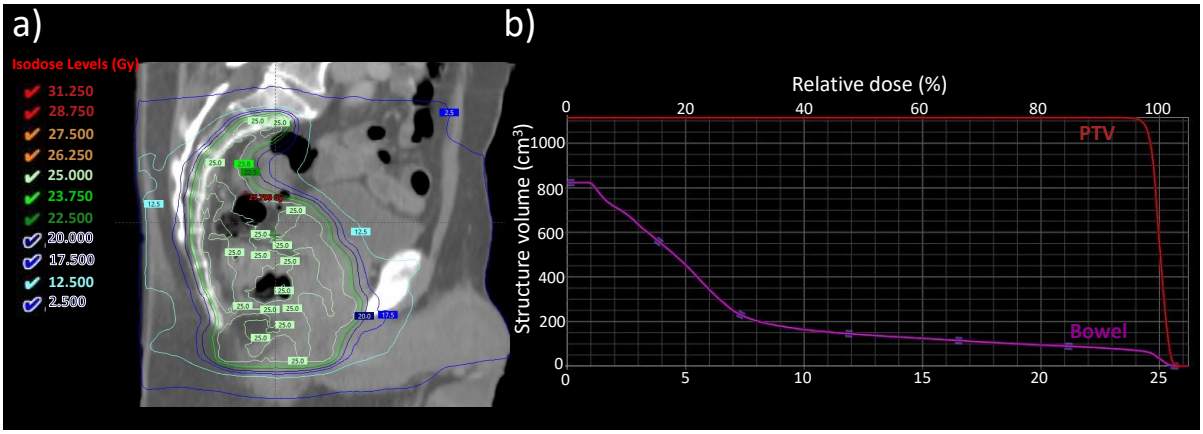

Figure 2.15: Isodose curves in a CT slice and dose volume histograms used in radiotherapy treatment planning.a) Sagittal plane of a synthetic CT (sCT) slice in a rectal cancer patient illustrating isodose curves within the pelvis for the isodose levels in Gy listed to the left. b) Cumulative Dose-Volume Histograms (DVHs) for the planning target volume (PTV) and the bowel for the rectal cancer patient. The DVHs give information about the absolute structure volume measured in cm^3 receiving at least an absolute dose measured in Gy. The relative dose is given above, where 100% relative dose means 100% of the prescribed dose which is 25 Gy.

2.6.3 Definition of treatment volumes

The ICRU Report 50 have defined different volumes that must be identified and delineated prior to RT [52]. The volumes add margins to the tumour and OARs to ensure sufficient target coverage and OAR sparing [6]. A simple illustration of the volumes are given in Figure 2.16. Gross tumour volume (GTV) is the visible or palpable tumour, clinical target volume (CTV) includes the tumour's subclinical spread, internal target volume (ITV) takes physiological motion and the CTV's possible change in size and position into account, and the PTV adds an internal margin (IM) and setup margin (SM) to the CTV in order to account for setup variations and patient motion [52]. OARs require protection, and planning organ at risk volume (PRV) can be defined to ensure low dose covering [52].

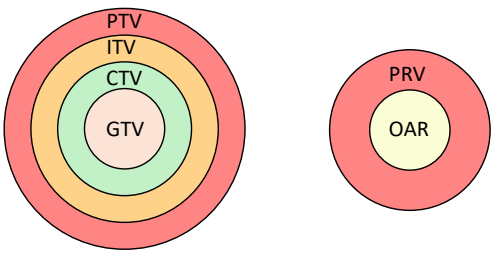

Figure 2.16: Simple illustration of the gross tumour volume (GTV), clinical target volume (CTV), internal target volume (ITV), planning target volume (PTV), organ at risk (OAR) and planning organ at risk volume (PRV).

In RT of rectal cancer, the Norwegian national guidelines for colorectal cancer pro-

vides guidelines for definition and delineation of the GTV and CTV. The GTV shall include the primary tumour and the bowel lumen, called GTVp, p indicating primary tumour. Other malignant tissue such as diseased lymph nodes are included in the GTVn. Elective CTV called CTVe_46 or CTVe_25, where the number represents a total radiation dose of 46 Gy or 25 Gy respectively, is delineated around the GTV, including relevant lymph node stations. The primary CTVp/n_50 is delineated around regions receiving a higher boost dose than the elective CTV [26]. Figure 2.17 illustrates a planning CT of a rectal cancer patient with segmented GTVp, CTVe_46, PTVe_46 (elective PTV receiving 46 Gy), CTVp_50 (primary CTV receiving a boost dose, the total dose equals 50 Gy) and PTVp_50 (primary PTV receiving a boost dose). The GTVp is located within the rectum, which is the target volume. As illustrated in Figure 2.17, other structures such as the bladder and the bowel also falls within the PTV and CTV, i.e., regions of high dose coverage.

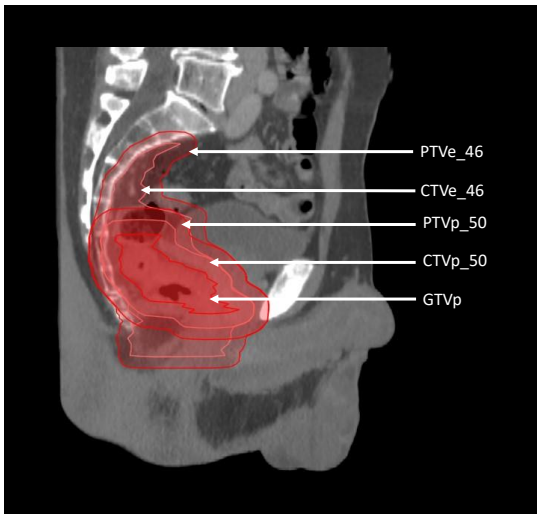

Figure 2.17: Illustration of a planning CT slice of a rectal cancer patient with segmented treatment volumes. The planning target volumes PTVe_46 and PTVp_50 are the elective PTV receiving 46 Gy and the primary PTV receiving 50 Gy respectively. The clinical target volumes CTVe_46 and CTVp_50 are the elective CTV receiving 46 Gy and the primary CTV receiving 50 Gy respectively. The gross tumour volume GTV_p indicates the primary tumour.

Error sources occurring during RT planning and the fractionated delivery limit the accuracy of the treatment [67]. During pelvic RT of rectal cancer, the small bowel and colon can change position and shape due to motion associated with digestion, giving rise to both intra- and interfractional uncertainties. The bowel may move during a fraction (intrafractional) and change its position relative to the pelvic bony anatomy from one fraction to another (interfractional). Interfractional uncertainties may also be caused by variations in the daily patient setup. Safety margins are required to ensure that the majority of the patients receive adequate dose coverage of their target volumes despite of uncertainties. Van Herk provided an overview of margin recipes accounting

for errors in RT [67]. Delineation inaccuracies, unknown extent of subclinical spread, organ motion and patient setup variations are the main sources introducing geometrical uncertainties and errors to the treatment, shifting the high dose regions away from the CTV. Systematic errors, such as delineation uncertainties, are important because they affect the tumour and OAR dose coverage identically at all treatment fractions. Random errors influence treatment fractions differently. Organ motion such as intra- and interfractional bowel motion can cause both systematic and random errors [67]. Systematic and random errors for an individual patient can not be determined until all treatment fractions are completed [68]. Errors measured in a population of similar patients treated in multiple RT fractions are analysed to determine the standard deviation (SD) σ for the random error, Σ for the systematic error, and the overall mean error for the entire population. Σ is the SD of the overall mean describing if the treatment preparation is performed in a reproducible manner and σ is the mean of the SD per patient determined after multiple treatment fractions [67].

2.7 Computed tomography and cone-beam computed tomography

CT and CBCT are imaging techniques acquiring 3D volumetric images of the patient [52]. While CT is used in both radiology and RT, CBCT is utilised in RT. A transversal slice from an MRI, CT and CBCT scan of a randomly selected patient included in this study are illustrated in Figure 2.18. The illustrated MRI and CT scans are acquired as part of the overall treatment planning, and the CBCT scan are acquired prior to the first treatment fraction. As illustrated, MRI offers better imaging of the soft tissue and higher contrast than CT, while CT better illustrates the bony anatomy. CT offers better contrast and resolution than CBCT.

Figure 2.18: One transversal slice from a) an MRI scan, b) a CT scan and c) a pre-CBCT scan of the pelvis of a rectal cancer patient.

2.7.1 Cone-Beam Computed Tomography

CBCT systems are mounted on the linac, as illustrated in Figure 2.19. Thereby, image data of the actual treatment conditions can be acquired, such as the actual position of the target volume and internal structures prior to or after RT delivery. In kilovoltage CBCT (kVCBCT), the radiation source and flat panel detector are attached to the linac, opposite to one another. An imaginary line between the radiation source and the detector panel is perpendicular to the direction of the therapeutic photon beam [52]. In megavoltage CBCT (MVCBCT), the radiation source is the therapeutic photon beam itself and the detector, an electronic portal imaging device (EPID), is placed opposite to the gantry. kVCBCT produces images of better resolution and contrast than those from MVCBCT and is the most utilised CBCT technique [52].

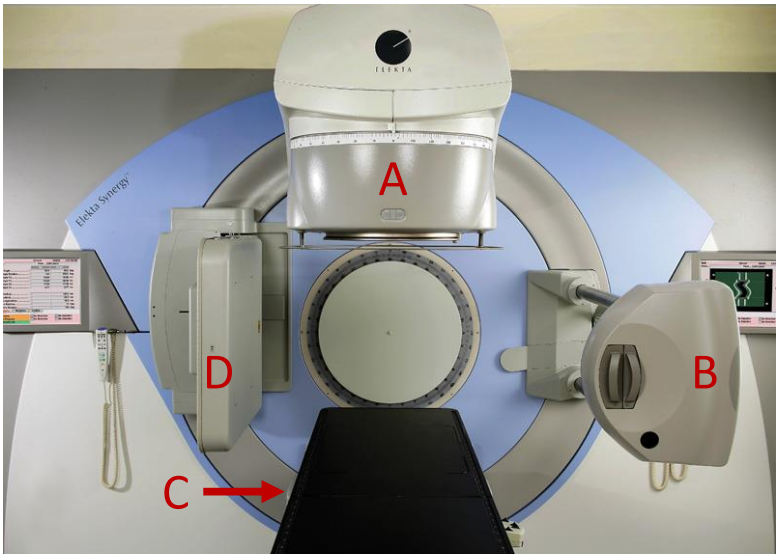

Figure 2.19: Illustration of an Elekta Synergy Linear Accelerator with a kV and MV imaging system. A: MV source. B: kV source. C: MV detector. D: kV detector. Illustration from [69], edited.

In CT and CBCT, photon beams travel through the patient's body before they hit radiation detectors placed on the opposite side of the radiation source, relative to the patient. The detectors measure attenuated photon intensity because the primary intensity of the radiation beam is attenuated on its way through the body due to photon interactions [52, 64]. In conventional CT, the radiation source, typically emitting radation beams of energies between 80 and 140 kilovoltage (kV), and the line detector are translated together in a circular rotation. Photon intensity attenuations are typically measured along the angular range of 360° of the circular motion in a fan-beam geometry for each angle in consecutive order, generating one transversal scan. Several such transversal scans imaged in 2D together form a 3D representation of the scanned body [52, 64]. In CBCT, however, the radiation detectors are inserted in a flat panel. The

photon source in CBCT emits photons in the kV or MV energy range towards the detectors, rotating at least 180° about the patient [52]. In contrast to conventional CT, the CBCT technique utilises cone beams while scanning the patient, as illustrated in Figure 2.20. Using a flat panel detector, information about the attenuations in different tissues are projected to different detector rows for different projection angles, not to detectors in one row only. Therefore, the CBCT system does not need to rotate around the patient multiple times to create 3D volumetric images. Hence, CBCT is less time-consuming than conventional CT. A disadvantage with CBCT, on the other hand, is the reduced image quality and images occupied by more noise and artifacts compared to CT images. The field of veiw (FOV) in CBCT is also smaller compared to conventional CT [64].

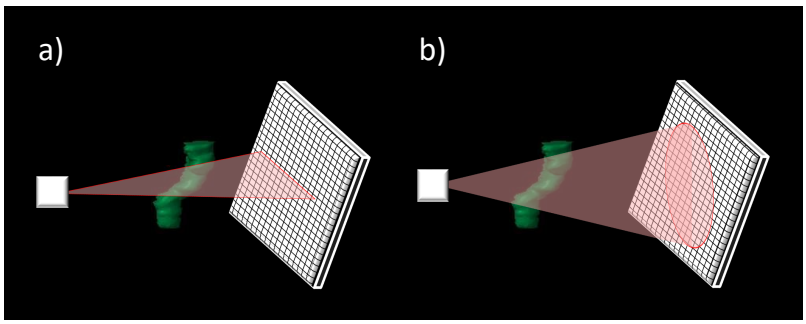

Figure 2.20: Illustration of a) fan beam in conventional CT to cover one slice of the illustrated rectum and b) cone beam and flat panel detector used in CBCT to cover multiple slices of the rectum such that one rotation can reconstruct a 3D volumetric image.

Common artifacts in CT and CBCT

Artifacts, meaning "artificially made", are artificial structures deviating from reality, created by the imaging system [64]. Artifacts are induced by disagreements between the mathematical assumptions of the reconstruction algorithms and physical conditions of the measuring set-up, often recognised as streaks and shadows in the images [70]. In the presence of artifacts, the CT and CBCT images may not reproduce and illustrate the actual anatomy of the body [64]. Images of reduced quality, occupied by artifacts, can influence the RT treatment planning process, resulting in inaccurate strucutre delineations and dose distribution calculations, consequently resulting in poor dose coverage of targets and/or overdosage of OARs [71]. Important artifact sources in CTs and CBCTs are patient movement, scattered radiation, metallic implants, and the patient exceeding the limits of the measurement field, among others [64]. CBCTs are normally more occupied by artifacts than CTs. For instance, artifacts due to the patient exceeding the limits of the measurement field occurs more frequently in CBCT than CT due to the reduced FOV in the CBCT scan [64]. One important source to arti-

facts in CBCTs is scattered radiation, which causes streak artifacts, reduces soft-tissue contrast and disturbs all density values. When photons travel from the radiation source towards the detector, they may be diffracted from their original path. Scattered photons will be detected in addition to the primary photons, resulting in an excessive number of photons being detected which causes an underestimation of the matter's ability to attenuate photons. A larger detector entails a higher probability of detecting scattered photons. Therefore, CBCTs are more affected by scattered photons than CTs, because of the larger flat panel detector compared to the line detector [70]. There are, however, methods to correct for different artifacts, improving the image quality [64].

A CT and CBCT slice occupied by metallic artifacts appearing as bright and dark streaks and shadows are illustrated in Figure 2.21 a) and b), respectively, where the patients have a metallic hip implants [71]. The reason why the artifact seem to be worse on the CT may be due to a possible difference in the metal type. The size and shape of the metallic implant as well as the atomic number and density of the metal impacts the degree of metal artifacts [71]. Figure 2.21 c) and d) illustrate artifacts in a CT and CBCT slice, respectively, caused by the large amount of air in the illustrated patients' bowels. The CBCT slice is more heavily affected by air artifacts compared to the CT slice.

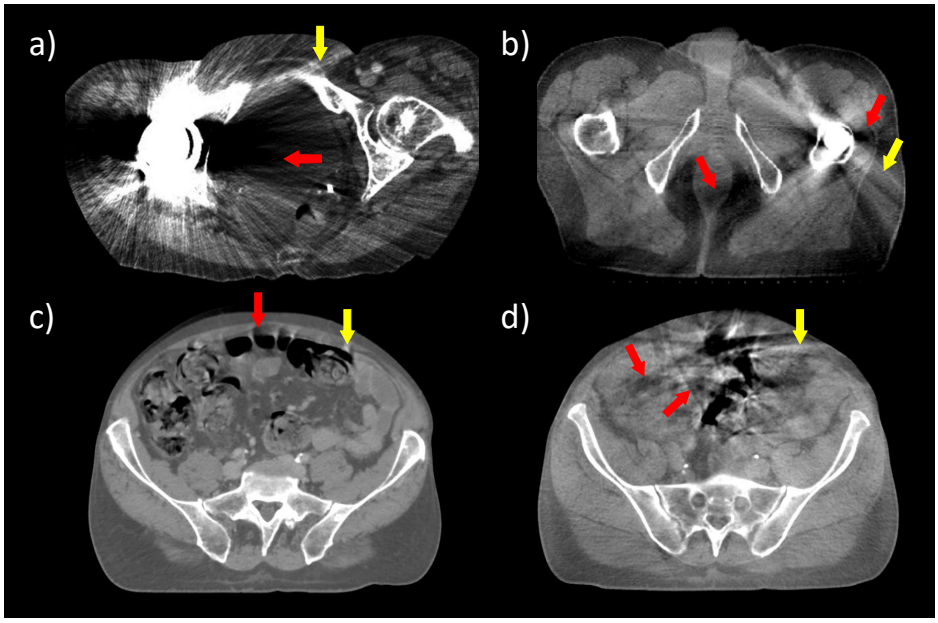

Figure 2.21: Illustration of artifacts in CT and CBCT slices. a) and b) illustrate metal artifacts in a CT slice due to a hip implant (illustration from [71], edited) and a CBCT slice due to a hip prosthesis (illustration from [72], edited) respectively. c) and d) illustrate air artifacts in a CT and CBCT slice respectively. Yellow arrows point at bright streaks and red arrows point at dark streaks/shadows.

2.8 Radiotherapy delivery techniques

2.8.1 Image guided radiation therapy

Image Guided Radiotherapy (IGRT) is an RT procedure where volumetric images of the patient's anatomy such as CT, MRI or commonly CBCT images are used to guide various steps of the RT process in order to ensure an accurate dose delivery. Imaging modalities are used to correct uncertainties related to inter- and intrafractional variations to ensure accurate target and OAR localisation [52]. The position of the RT targets and OARs are examined just before, during and immediately after a treatment fraction.

The kV imaging system integrated on modern linacs (Figure 2.19) can acquire planar 2D images in addition to CBCT images. The planar images are used to verify the patient set-up relative to the treatment room coordinate system to ensure consistent patient positioning in all fractions, and to determine where the planned targets are positioned relative to for instance bony landmarks, which are landmarks in the patient's fixed bony structure. Fiducial markers can be inserted into target structures in for instance the pelvis, and any motion of the fiducial markers and hence the targets can be observed and tracked during RT delivery [52].

IGRT is performed at different frequencies, either weekly or daily. Daily IGRT is performed at each RT fraction to observe and correct for systematic and random errors. Weekly IGRT only corrects for systematic errors present in all fractions of the week [73]. Planning CTs are acquired as part of the treatment planning process. It is important that the patient is positioned in the same way during planning and treatment delivery. Immobilisation devices may be used both during the CT scanning and treatment delivery to ensure that the patient remains in a fixed position, increasing the accuracy and reproducibility of the RT delivery. When new images are acquired, such as planar kV images or CBCTs acquired in the treatment room in daily IGRT prior to each fraction, the new images are related to the planning CT. Image registration is carried out where the image of the day is correlated to and fused with the planning CT based on mapping of corresponding structures such as the patient's bony anatomy. Based on the image fusion, the treatment table where the patient is placed is translated such that the image of the day matches and aligns with the planning CT [52]. Images can also be acquired immediately after a fraction is delivered to observe if the target was located within the defined PTV and high dose regions during delivery, or to observe any other location shift of internal structures.

2.8.2 Adaptive radiotherapy

ART is an IGRT based treatment technique, first introduced in 1997 by Yan et al. as "a closed-loop radiation treatment process where the treatment plan can be modified using a systematic feedback of measurements" [74]. In ART, the treatment plan is altered during the RT course to account for changes in the radiation target and normal tissue to better fit the patient's anatomy and physiology of the treatment day. Targets and structures may change size, shape, functionality and dose response through the entire RT course. The aim of ART is to improve the clinical outcomes by accurately deliver the dose to the targets while minimising dose delivered to the normal tissue. Since the plans are adapted to the patient's internal anatomy of the treatment day, the margins around the target and the OARs can be reduced. ITVs can be removed or shrinked and therefore smaller volumes are irradiated if the intrafractional motion is assumed to be small and the patient's anatomy during RT delivery is the same as during replanning. ART requires volumetric images such as CBCTs or MRIs to create a new dose plan or to modify an existing one. The volumetric images are used to detect internal anatomical or physiological changes, and a choice is made whether to adapt the treatment plan or to keep the original plan. This assessment can be a manual or a complex automatic evaluation, including structure segmentations and recalculation of the dose, among others. The need of an adapted plan must be evaluated. If the adapted plan has too few changes compared to the initial plan, producing no dosimetric improvements, or if it violates predefined OAR or target constraints, it can be rejected [6].

ART can either be applied offline where the replanning is carried out between fractions, online immediately before the treatment fraction is delivered, or inline during a fraction. In replanning, a standard TPS can be utilised for offline ART, while specialised systems are required for online and inline ART. Offline ART is used to account for functional changes in the tumour or normal tissue. Images acquired after the previous treatment fraction are used to replan the next fraction. However, if the internal changes happen quickly such as intrafractional motion or greater interfractional variations, offline ART may not be able to adapt to these changes. Online ART can, among others, be useful in pelvic RT where the bowel is located, to account for the rapid internal changes of the bowel's position due to the constant peristalsis waves of the bowel. The treatment plan is immediately adapted to changes observed in the images and the aim of online ART is to deliver the treatment before further major changes can happen. However, online ART struggles to adapt to rapid and unpredictable intrafractional changes, such as bladder filling, which can occur between the time of imaging and RT delivery or during the delivery [6]. The entire online ART workflow is per-

formed while the patient remains in the treatment position at the treatment table, which may be uncomfortable for some patients. Online ART prolongs the treatment duration compared to conventional IGRT. Therefore, online ART requires more specialised tools than offline ART to ensure that the process does not become too time-consuming [6]. Ethos™Therapy [5] and MRI Linacs are expensive systems dedicated to conduct online ART. The systems acquire daily volumetric images used to create a daily and unique treatment plan based on the patient's anatomy of the day [6]. Using AI and machine learning (ML), the Ethos™Therapy, which is a Varian Adaptive Intelligence™solution, creates contours of the daily imaged OARs and target structures and generates adaptive treatment plans while the patient is on the treatment table [5]. DVHs based on Ethos' segmentations are generated while the auto-segmentations are manually controlled and possibly edited if they do not sufficiently fit the structures. If the segmentations are edited, the adaptive plan must be replanned and new dose distributions calculated by the system, resulting in an extended duration of the treatment time [6]. Typically, one treatment using Ethos™Therapy has a duration of 15 min, including plan selection and RT delivery [5]. Online ART without such dedicates systems are more time-consuming. A plan library approach can be utilised to deliver online ART if dedicated systems are not available. A plan library consists of dose plans created in advance of the adaptation, based on predictable potential anatomical changes such as bladder filling. The plan from the library that best matches the patient's anatomy-of-the-day is selected during treatment [6]. Figure 2.22 shows a schematic illustration of an online ART workflow using Ethos™Therapy and CBCT as imaging modality. Post-CBCT is a part of the workflow, used to observe if the targets and OARs have moved during treatment.

Figure 2.22: Schematic illustration of an online adaptive radiotherapy (ART) workflow using Ethos™Therapy. Cone-Beam Computed Tomography (CBCT) are used to acquire volumetric images before (pre) and after (post) the radiotherapy (RT) is delivered. Tasks in orange boxes are performed by the treatment personnel. Tasks in blue boxes are performed by systems dedicated to conduct online ART, such as Ethos™Therapy, using artificial intelligence (AI) to delineate structures in the CBCTs and generate ART plans.

De Jong et al. showed that online ART using a plan library approach and dedicated systems can be useful in rectal cancer RT to achieve improved target coverage and OAR sparing [75, 76]. They developed a library of plans with variable margins to the upper anterior mesorectum, since changes in the rectal volume are the main reason why the target (rectum) volume changes shape, and the upper mesorectum has the

largest deformations [75]. Creating the adaptive plan library prolonged the treatment planning process, and the total ART workflow lasted 120 min longer than a non-ART workflow [75]. Daily pre-CBCT scans were acquired before each fraction, and the plan with the smallest PTV covering the entire CTV on the current CBCT was selected [75]. They found that the ART plan resulted in an increased volume of the mesorectum receiving at least 95% (V95%) of prescribed dose compared to the non-ART plan [75]. Figure 2.23 illustrates the differences of the small bowel volume (an OAR) receiving dose in ART and non-ART [75]. As the figure illustrates, ART significantly reduced the small bowel V15, V40, V45 and V95% compared to non-ART for the patients receiving LCRT [75]. In 2021, de Jong et al. described the first clinical experience of CBCT based online ART for neoadjuvant treatment of rectal cancer [76]. Their investigation included rectal cancer patients scheduled for SCRT of 5×5 Gy delivered using Ethos™Therapy [76]. The entire online ART workflow had a duration of 26 min on average for all patients, with no interrupted fractions [76]. For all fractions, the PTV V95% requirement was almost always met by the adaptive plan (except in 3 of 60 fractions), but not by the scheduled plan where the PTV V95% was less than the requirement in 55 of 60 fractions [76]. During treatment, the adaptive plan was always chosen [76].

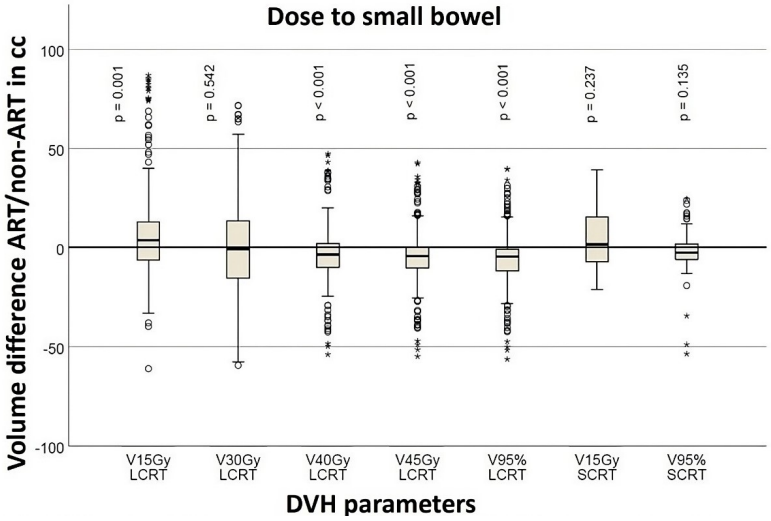

Figure 2.23: Box plots illustrating volume differences for different small bowel dose-volume parameters between adaptive treatment (ART) plans and a non-adaptive treatment plan (non-ART) for long and short course radiotherapy (LCRT and SCRT, respectively) by de Jong et al. [75]. $p < 0.05$ indicate statistically significant volume differences. A negative volume difference implies that the adaptive plan resulted in less dose to the small bowel than the non-adaptive plan. Illustration from [75], enhanced.

2.8.3 Segmentation

Standard definitions of delineation techniques for normal tissue structures are important in order to compare findings from clinical trials and improve and evaluate normal tissue dose constraints [77]. A Radiation Therapy Oncology Group (RTOG) atlas provides male and female pelvic normal tissue contouring guidelines [77]. The contouring atlas contains, among others, delineation consensus definitions of the small bowel, colon, bowel bag and rectum, describing how the structures should be contoured based on their shape, extent or location relative to other internal structures visible in CT images [77]. The bowel bag offers a faster way to contour the bowel, however, the atlas does not suggest which bowel contouring strategy are most appropriate in clinical situations [77]. A RTOG contouring consensus atlas for the elective CTVs to be used in the planning of pelvic IMRT for anal and rectal cancers is also developed [78].

Manual segmentation of RT targets and OARs is a comprehensice and time consuming task. Accurate structure segmentations during RT planning can increase the tumour control and reduce the risk of toxicities [10]. Manual segmentations can introduce uncertainties due to intra- and inter-delineator variations. ML algorithms are developed to perform accurate and efficient auto-segmentations of anatomical structures and to improve the consistency in normal tissue segmentations in medical images. Auto-segmentation algorithms are useful if their segmentation performances are better or comparable to manual segmentations while reducing the clinicians' workload. In recent years, deep learning algorithms have become the dominant commercial offer of auto-segmentation algorithms, with multiple available clinically useful software packages [10]. For instance, Varian provides Ethos™Therapy which is a Varian Adaptive Intelligence™solution using AI and ML to automatically segment structures [5]. The clinical use of auto-segmentation methods are limited by, among others, the lack of standard contouring protocols and trust among users. Thus, in RT planning, auto-segmentation algorithms are normally manually controlled and edited by clinicians [10].

Auto-segmentation methods

An artificial neural network, illustrated in Figure 2.24, is an ML model (an application of AI) consisting of connected nodes organised in one or multiple layers. The input layer receives input data which are propagated through the network towards the output layer. Calculations are performed at every neuron, resulting in an output value. Layers between the input and output layer are called hidden layers. The structure of the artificial neural network of an ML model can be expressed mathematically, characterised by

parameters and hyperparameters that are optimised to improve the performance of the algorithm. Artificial neural networks can have multiple hidden layers, a hyperparameter, increasing the complexity of the network. Deep learning refers to artificial neural networks with a large number of hidden layers [10].

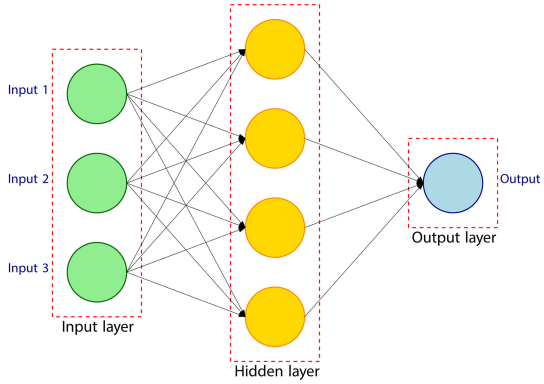

Figure 2.24: Illustration of an artificial neural network with an input layer accepting three input values (green neurons), one hidden layer (yellow neurons) and an output layer (blue neuron). All neurons are connected. Input values are propagated through the network where calculations are performed to determine an output value [10]. Illustration from [10].

There are multiple different methods for auto-segmentation of medical images of various complexities, including AI methods. Examples of simpler auto-segmentation methods of CTs are intensity analysis and thresholding, where a voxel, which is a component of a CT image, is assigned to a structure if its intensity measured in Hounsfield Units (HU) is above a specified threshold value, or based on their neighboring voxels' intensity values. Atlas-based auto-segmentation algorithms are other relative simple algorithms, which segment structures based on anatomical information. A manually segmented reference image, called an atlas, are registered to the studied image and segmented structures from the atlas are mapped to the studied image. Deep learning segmentation algorithms aim to detect a mathematical model to predict an output label based on an input, and hence identify anatomical structures in images. The convolutional neural network (CNN) model or the U-net model, which is a design of the CNN architecture, are deep learning models, frequently used to segment and identify structures in medical images [10].

2.8.4 Evaluation of segmentation performance

Multiple metrics are available to evaluate the geometric accuracy of segmentations [79]. Segmentation performances can also be evaluated dosimetrically [80].

Geometric evaluations

Geometric evaluation is the most common method for evaluating segmentation performances. Examples of metrics used in evaluating segmentation performances of medical images are the Dice Similarity Coefficient (DSC), the Hausdorff Distance (HD) and the centroid distance, among others. These are overlap, spatial distance, and moment based evaluation metrics, respectively [10, 79].

The DSC is commonly used in comparing medical volume segmentations [79]. It was first introduced by Lee R. Dice in 1945 as a "Coincidence Index" with the purpose of measuring the association between species [81]. DSC measures the similarity between shapes such as two structure segmentations by comparing their overlap [79]. The DSC can take values in the range [0,1]. If the DSC equals zero, there are no overlap between the compared volumes. If there are complete overlap between volumes, the DSC equals one, which is the ideally case when comparing different segmentations of the same structure. Hence, the DSC between two segmentations are easy to interpret. Therefore, and because the DSC between two structure segmentations is easy to compute, it is frequently used in evaluating segmentations in medical images [79, 80]. When testing and evaluating auto-segmentation algorithms, as demonstrated by for example Boehringer et al. in their study on deep learning algorithms for tumour segmentation from brain MRI images, a DSC of 0.80 is often used as a threshold to signify acceptable segmentation performances [82].

The HD is used to measure the closest distance between the outer borders of the compared structure segmentations. There are several variants of the HD, such as the average HD or the 95^{th} percentile HD (HD95) [79, 80]. Since HD is sensitive to noise and outliers, it was proposed to use a quantile method, less sensitive to outliers [79, 83]. HD95 represents the 95^{th} percentile of the distances calculated between the points of two structure segmentations [79]. A greater dissimilarity between segmented structures is represented by a higher HD value. Ideally, the HD, or HD95 value, should equal zero when comparing different segmentations of the same structure, meaning that there are no discrepancies between the segmentations.

The centroid distance is used to compare the center of mass between two segmentations. It is suitable to evaluate segmentation performances of small structures of approximately spherical shapes. However, the centroid distance is less suitable to evaluate structure segmentations of complex shapes [84].

Dosimetric evaluation

Dosimetric calculations can be used to investigate the dose difference between different segmentations of the same structure and thus the clinical impact of the segmentations.

Geometric segmentation variations may not necessarily indicate significant differences in delivered dose. For instance, geometric metrics such as DSCs will result in good agreement for two nearly overlapping segmentations. However, the small segmentation difference can be of clinical importance if the structures are located within regions where the amount of deposited dose changes rapidly. Similarly, segmentations of poor geometric agreement located in regions of constant dose coverage may have high dosimetric agreement, meaning no clinical relevance [80].

2.9 Bowel as a risk organ in radiotherapy

The bowel is an OAR in RT of rectal cancer, which may develop acute or chronic side effects if exposed to radiation [1]. Hence, accurate definition of the bowel during online ART planning is important to reduce bowel exposure. Uncertainties regarding the dose deposited in the bowel during pelvic RT are induced when bowel motions are not accounted for, since the bowel's actual position during RT delivery then are unknown. In addition to the bowel moving itself, changes in the bowel's position is affected by changes in size and shape of other nearby organs, such as the bladder [85, 86].

2.9.1 Bowel anatomy

The digestive system breaks down food and beverages, absorbs nutrients, and carries waste products out of the body in the form of stools. It consists of the GI tract in addition to the liver, pancreas and gallbladder. The GI tract consists of multiple organs from the mouth to the anus, one of them is the bowel. The bowel, also known as the intestine, is further categorised into the small and large bowel [87]. Figure 2.25 illustrates the anatomy of the lower digestive system including the small bowel, colon (large bowel), rectum and anus.

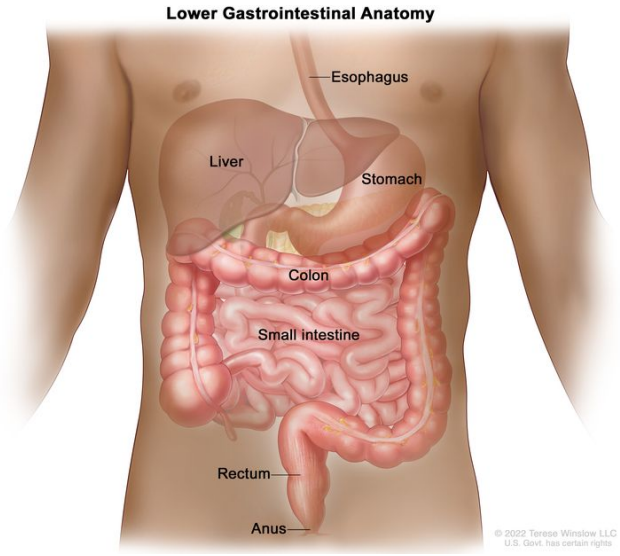

Figure 2.25: Anatomy of the lower digestive (gastrointestinal) system. Illustrated parts of the gastrointestinal tract are the esophagus, stomach, small bowel (small intestine), large bowel (colon), rectum and anus. The liver is also a digestive organ. Illustration from [13].

The small bowel consists of three parts; duodenum, jejunum and ileum. The jejunum and ileum are attached to the mesentery. The mesentery has a fan like shape and is attached to the posterior wall of the abdominal cavity. Since these loops are attached to the mesentery and not directly to the abdominal wall, they are quite mobile and can change their position due to the peristaltic movements [88].

The large bowel consists of the caecum, vermiform appendix, colon and rectum. The colon can be divided into an ascending, transverse and descending part and the sigmoid. It surrounds the loops of the small bowel [89, p. 226]. The ascending colon is connected to the small bowel, with the transition located approximately 6-7 cm above the most inferior part of the ascending colon. The descending colon passes into the sigmoid which is connected to the rectum where the stool is eliminated [57].

While the small bowel and colon are located in the abdomen, the rectum is located in the pelvis [57].

2.9.2 Bowel structure definition in radiotherapy planning

The bowel bag, peritoneal space or intestinal cavity technique is frequently utilised instead of the bowel loop technique when delineating the bowels in CTs or CBCTs in treatment planning of RT. Several studies have shown that delineating the bowel bag structure, or adding margins to the bowel loops, in both pre- and postoperative rectal, prostate and gynecologic cancer patients can replace the individual bowel loops to spare the bowel for dose exposure in non-adaptive IMRT treatments [7, 90, 91]. There are advantages and disadvantages to both delineation methods. Segmenting the bowel bag

in CTs or CBCTs reduces the workload compared to segmenting each individual bowel loop because the volume is easier to define and faster to delineate [92]. The bowel bag encompasses the entire bowel, despite of bowel motion, and thus accounts for all regions where the small bowel or colon can be located [7]. Limiting the dose to the bowel bag ensures that the dose to the bowel is limited for all bowel segments, and increases bowel protection when the bowel motion otherwise is not accounted for. However, dose constraints for the bowel bag can reduce the dose coverage to the PTV to a greater extent than dose constraints for the individual bowel loops. Applying dose constraints for the individual bowel loops less restricts the dose that can be deposited in the target volume. Then, it is crucial to know where the bowel is located during treatment to avoid overdosage of bowel segments.

2.10 Bowel dose response

Radiation induced damages to tissues are classified as stochastic or deterministic side effects. A stochastic side effect has severity independent of the dose received by the tissue, but the probability of the side effect to occur increases with increasing dose. A deterministic effect has increasing severity for increasing dose above a threshold [53]. Normal tissue can develop acute or late side effects as a response to radiation. Various tissues respond differently to radiation and are characterized by different α/β values (see Section 2.4.2). Performing regression analysis on clinical data, relating dose distributions to certain toxicity endpoints, is a common method to generate dose-response models for different tissues [93]. The Common Terminology Criteria for Adverse Events (CTCAE) provides a grading scale to determine the severity of adverse events (AEs) used as toxicity endpoints in dose-response modelling. According to the CTCAE version 4, an AE is a negative sign, such as a disease, which potentially can be caused by a medical treatment or procedure [94]. The severity of an AE is graded in grades 1 through 5, ranging from mild to life-threatening symptoms [94]. Figure 2.26 shows an overview of the clinical descriptions of the severity of diarrhoea, an AE, of grades 1 through 5.

Adverse Event	1	2	3	4	5
Diarrhea	Increase of <4 stools per day over baseline; mild increase in ostomy output compared to baseline	Increase of 4 - 6 stools per day over baseline; moderate increase in ostomy output compared to baseline	Increase of >=7 stools per day over baseline; incontinence; hospitalization indicated; severe increase in ostomy output compared to baseline; limiting self care ADL	Life-threatening consequences; urgent intervention indicated	Death
Definition: A disorder characterized by frequent and watery bowel movements.					

Figure 2.26: Clinical description of the severity of the adverse event diarrhoea of grades 1 through 5. Semi-colon (;) refers to "or". ADL: Activities of Daily Living. Image from [94].

The reliability of dose-response models are weakened when the dose distributions only are estimates of the dose delivered to the tissues, influenced by both segmentation uncertainties and organ motion [93]. Different bowel delineation techniques, such as the bowel bag or bowel loop technique, contributes to increased uncertainties in dose-response models for the bowel when the studied bowel volumes are inconsistent.

Baglan et al. quantified dose-volume relations between irradiated small bowel loops and the grade of acute small bowel toxicity during LCCRT of rectal cancer using a three-field technique [95]. They showed that both low and high dose-volumes were significantly associated with grade ≥ 3 acute small bowel toxicity [95]. Also, they discovered that the small bowel volume exposed to at least 15 Gy (V15) was highly associated with grade ≥ 3 acute small bowel toxicity [95]. These findings were later tested and validated by Robertson et al. for an increased number of patients [96]. Hence the Baglan-Robertson threshold model illustrated in Figure 2.27 a), relating small bowel dose-volumes to risks of grade ≥ 3 acute small bowel toxicity based on the threshold value of V15 < 120 cc [1]. According to this threshold model, small bowel dose-volumes below the illustrated curve have an estimated risk of approximately 10% for grade ≥ 3 acute small bowel toxicity, while bowels above the curve have an estimated risk of approximately 40% for the same toxicity grade [1]. Roeske et al. studied the bowel similar to the peritoneal space structure, and showed that the volume of this bowel structure exposed to 45 Gy was the most statistically significant factor related to acute GI toxicity in pelvic IMRT [97]. This resulted in a threshold model where the V45 of this bowel volume should be less than 195 cc to reduce the risk of acute GI toxicity [97]. Based on the findings by, among others, Baglan et al., Robertson et al. and Roeske et al., QUANTEC later defined bowel dose constraints to reduce the risks of grade ≥ 3 acute small bowel toxicity by 10% [98]. V15 should be less than 120 cc when the individual small bowel loops are defined [1]. For bowel bag delineation, V45 should be less than 195 cc [1]. Dose-volume thresholds are believed to reduce the risk of late toxicities as well but have not yet been established [1].

Banerjee et al. investigated whether the peritoneal space contouring technique could be used instead of the bowel loop technique and still be predictable of grade ≥ 3 acute small bowel toxicity [92]. They showed that delineating the small bowel loops resulted in the most precise toxicity risk estimates at each 5-Gy interval between 5 and 45 Gy [92]. However, regardless of defining the bowel as individual loops or the peritoneal space, they found that both low and high doses were predictable of toxicity [92]. Hence, their findings suggest that there are no absolute threshold for acute toxicity inducing doses, such as the Baglan-Robertson threshold model or the model provided by Roeske et al. [96, 97]. Holyoake et al. recently (2019) performed a meta-analysis of

the relation between small bowel dose-volume parameters and acute toxicity in conventionally fractionated RT of rectal cancer [4]. Patients with grade ≥ 3 acute small bowel toxicity were shown to have significantly greater irradiated small bowel volumes for every 5-Gy interval between 5 and 40 Gy compared to patients experiencing toxicity of lower grades [4]. Hence, similar to Banerjee et al., they also found a continuous relationship between small bowel dose-volumes and toxicity, supporting the absence of a threshold dose that induces toxicity [4]. Both Banerjee et al. and Holyoake et al. estimated toxicity risks based on small bowel loops from planning CTs [4, 92]. Due to bowel motion, the bowels might be positioned in different locations during treatment then when the planning CTs were acquired. Therefore, their small bowel or peritoneal space dose-volumes may not be representative for the actual volumes exposed to doses, possibly explaining their findings of significant relationships between small bowel dose-volumes and toxicity for both low and high dose levels. Holyoake et al. found small bowel V10 to be the most statistically significant predictor of toxicity risks [4, 92]. Figure 2.27 b) illustrates the logistic regression curve provided by Holyoake et al. for toxicity probability estimation for small bowel loop V10 [4]. Figure 2.27 b) illustrates a wide distribution of the source data points (from the different studies included in their meta-analysis), each relating toxicity risk to small bowel V10, together with a wide 69% confidence interval [4]. The spread of the data highlights the uncertainties related to dose-response models for the bowel [4]. Holyoake et al. discussed multiple limitations to their study, reducing the accuracy of their dose-response model [4]. For instance, the studies included in their meta-analysis showed inconsistency in the utilised toxicity endpoints [4]. Also, the utilised fractionation schemes varied to some extent [4].

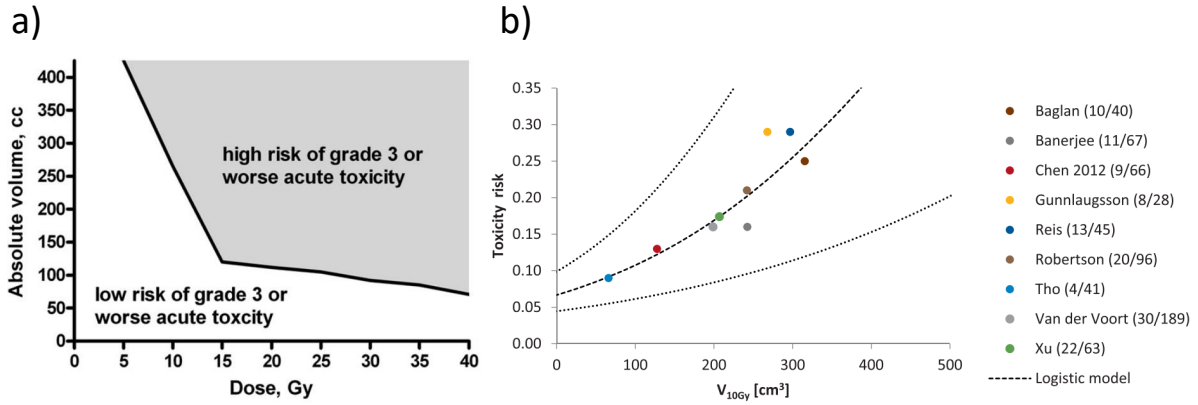

Figure 2.27: a) The Baglan-Robertson model, estimating the risk of grade ≥ 3 acute small bowel toxicity based on the dose-volume relationship of irradiated individual small bowel loops. Illustration from [1]. b) Reported toxicity plotted against small bowel V10 from each study included in the meta-analysis by Holyoake et al., listed to the right. The logistic regression model are plotted in the dashed, bold line and the dotted lines indicate the 69% confidence interval. Illustration from [4].

Chapter 3

Materials & Methods

3.1 Patients

This master study includes 15 patients treated for rectal cancer with neoadjuvant concurrent chemotherapy and LCRT or SCRT. The research project was pre-approved by the Regional committees for medical and health research ethics (REK - 205367) and all patients were anonymized for this project. Before start of RT the patients were advised to avoid the consumption of food that causes a great amount of air in the GI tract, but they did not follow any fasting or hydration regimen before treatment. For each of the first five fractions a CBCT scan was acquired before (pre-CBCT) and after (post-CBCT) RT delivery. This investigation includes data from the first (fx1) and last (fx5) treatment fraction, except from two patients (ER2 and ER14) where the post-CBCTs of fraction one were missing and therefore fraction two was utilised instead. All scans were acquired in head-first supine position with a custom pillow for knee- and ankle fixation. The average (range) time in-between CBCT images was 7 min (5 min-11 min). The pre- and post-CBCTs were rigidly aligned based on the pelvic bone anatomy using six degrees of freedom. An overview is given in Table 3.1.

3.2 Segmentation

Anatomical structures of the pelvis were segmented on all pre- and post-CBCTs. The structures of interest were the bladder, small bowel and colon, since they are the main organs at risk in RT of rectal cancer, in addition to the rectum. The segmented small bowel and colon consisted of individual loops when possible. The small bowel structure did not distinguish between the duodenum, jejunum and ileum. The colon structure consisted of the ascending and descending colon and the sigmoid. The transverse colon was in all patients located more cranial than the most cranial CBCT slice, and was

Table 3.1: Overview of the gender of the patients, which radiotherapy course they received and the time in minutes between their pre- and post-CBCTs were acquired for fraction one and five. LCRT: long-course radiotherapy. SCRT: short-course radiotherapy. ER: EthosRecti (rectal cancer patient treated with Ethos™Therapy).

Patient Details			Time (min:sec) between pre- and post-CBCTs	
Patient	**Gender**	**Treatment**	**Fraction one**	**Fraction five**
ER1	Male	SCRT	07:31	07:06
ER2	Male	SCRT	06:42	07:55
ER3	Male	LCRT	06:54	05:40
ER4	Female	LCRT	05:23	04:38
ER5	Female	LCRT	05:51	06:59
ER6	Male	LCRT	07:09	05:49
ER7	Male	LCRT	06:25	05:46
ER8	Female	SCRT	08:19	08:04
ER9	Female	SCRT	06:26	07:22
ER10	Male	SCRT	08:10	07:57
ER11	Male	SCRT	10:01	07:57
ER12	Male	LCRT	06:19	05:09
ER13	Male	SCRT	10:50	07:41
ER14	Female	SCRT	07:09	-
ER15	Male	LCRT	08:00	06:05

therefore not included. The rectum was segmented to better distinguish between rectum, which is the radiation target, and bowel. The bladder segmentations were utilised to investigate intra- and interfractional changes in bladder volume.

3.2.1 Segmentation training

In order to learn how to distinguish between different anatomical structures in CT and CBCT images and to navigate in the body through transversal, sagittal and frontal planes, the candidate, referred to as delineator one, was trained by supervisor Johanna Austrheim Hundvin, referred to as delineator two. The CT scan of patient ER6 was initially used for learning to segment structures. The bowel was separated into bowel (as the small bowel and colon combined) and the rectum, according to the contouring guidelines provided by RTOG [77].

Delineator two guided delineator one in bowel segementation in CT images, followed by independent segmentation by delineator one. In case of doubt, or after finalised segmentation, delineator one compared the segmentations with existing segmentations (hidden during the delineation process) performed by delineator two. If the segmentations were not consistent, the delineators discussed the differences and revised the structures if necessary. Completed training on CT image quality was followed by delineation in CBCT images.

3.2.2 Manual segmentations

Delineator one and two manually segmented the structures of interest in the pre- and post-CBCTs following the consensus contouring guidelines of male and female pelvis in RT provided by RTOG [77]. An effort was made to segment the loops of the small bowel and colon, and not the bowel bag or peritoneal space. All structures of interest were segmented in fx1, while only the bladder and bowel were segmented in fx5 due to time limitations. Exceptions were made for patient ER2, ER3 and ER13, where, due to artifacts, it was decided not to distinguish between the small bowel and colon in fx1 but to segment the combined bowel. In both fractions, the segmented bowel structure in these patients became more similar to a bowel bag structure.

Delineator one segmented structures in each pre- and post-CBCT of fx1 and each pre-CBCT of fx5. Automatic segmentations performed by an AI algorithm, TotalSegmentator (TS) [99, 100, 101], were used as a starting point which delineator one corrected to better fit the structures. Mostly, delineator one first segmented the structures in the pre- and post-CBCTs of the first fraction for all 15 patients, before segmenting fx5. Delineator two segmented structures in all post-CBCTs.

Segmentation technique, fraction one

In each CBCT scan, delineator one started by segmenting the rectum. The rectum in the planning CT scan was already segmented as a part of the clinical treatment. When segmenting the rectum in pre-CBCTs, this pre-segmented CT rectum structure was copied to the pre-CBCT image and then corrected to fit the shape and location of the rectum of the pre-CBCT. When segmenting the rectum in post-CBCTs, the segmented rectum structure from the corresponding pre-CBCT was copied and corrected to fit the rectum of the post-CBCT. RTOG's contouring guidlines [77] were utilised to locate the transition between the rectum and the colon and the inferior extent of the rectum. The rectum ends superiorly when it loses its round shape, studied in the transversal plane [77]. It connects with the sigmoid which often has a more elongated shape in the transversal plane due to the bowel's anatomy, as illustrated in Figure 2.25. The segmentation of the rectum ends inferiorly at the lowest level of the right or left ischial tuberosity [77], the sitting bone. After completing the segmentation of the rectum in a CBCT, the colon was segmented. The colon starts inferiorly from where the rectum ends superiorly [77], and was segmented to the most superior slice of the CBCT scan. The small bowel was segmented after completing the colon segmentation of a CBCT scan. The small bowel was segmented inferiorly from the most inferior small bowel loop visible in the CBCT scan [77] and superiorly to the last CBCT slice of the scan.

Further, the bladder was segmented in each CBCT. In the pre-CBCTs, delineator one corrected TS's bladders, while in the post-CBCTs, delineator one copied delineator two's bladders and occasionally made some small adjustments to them.

When segmenting the rectum, colon, small bowel and bladder, the segmentations were mostly performed in the transversal plane. The sagittal and frontal planes were utilised to study whether the segmentations made a continuous colon or small bowel structure and where it would make sense that the structures were located within the patient. The inbuilt function "crop structure" in Eclipse was used to ensure that the segmented structures never extend into one another. A bowel structure, which is small bowel and colon combined, was created as a union of the segmented small bowel and colon structures.

Segmentation technique, fraction five

Due to the time spent on delineation in fx1 where one scan took approximately 4-6 hours, a choice was made to segment only the combined bowel in the CBCTs of fx5, without first distinguishing between small bowel and colon as in fx1. In addition to the bowel, the bladder was segmented in each CBCT of fx5. The segmentations in fx5 were carried out in a similar manner as explained for fx1, mostly in the transversal plane, from the most inferior to the most superior CBCT slice containing the studied structures.

3.2.3 AI segmentation

For each patient, one CT scan, five pre-CBCTs and five post-CBCTs were exported as DICOM files from Eclipse and made available for an AI algorithm for automatic segmentation, TotalSegmentator version one (v1) [99, 100, 101]. TS is an nnU-Net segmentation algorithm, based on the U-Net architecture, trained on different CT images to automatically segment major anatomical structures of the body [101]. The algorithm was run on all the CT and CBCT scans, and then all the files with their corresponding AI segmentations were imported back to Eclipse. Figure 3.1 shows a CBCT scan of the pelvic region in one patient, including segmented structures by TS and a list of all structures the algorithm delineated in this patient. TS took about 15-20 min to segment all these structures for one CBCT scan.

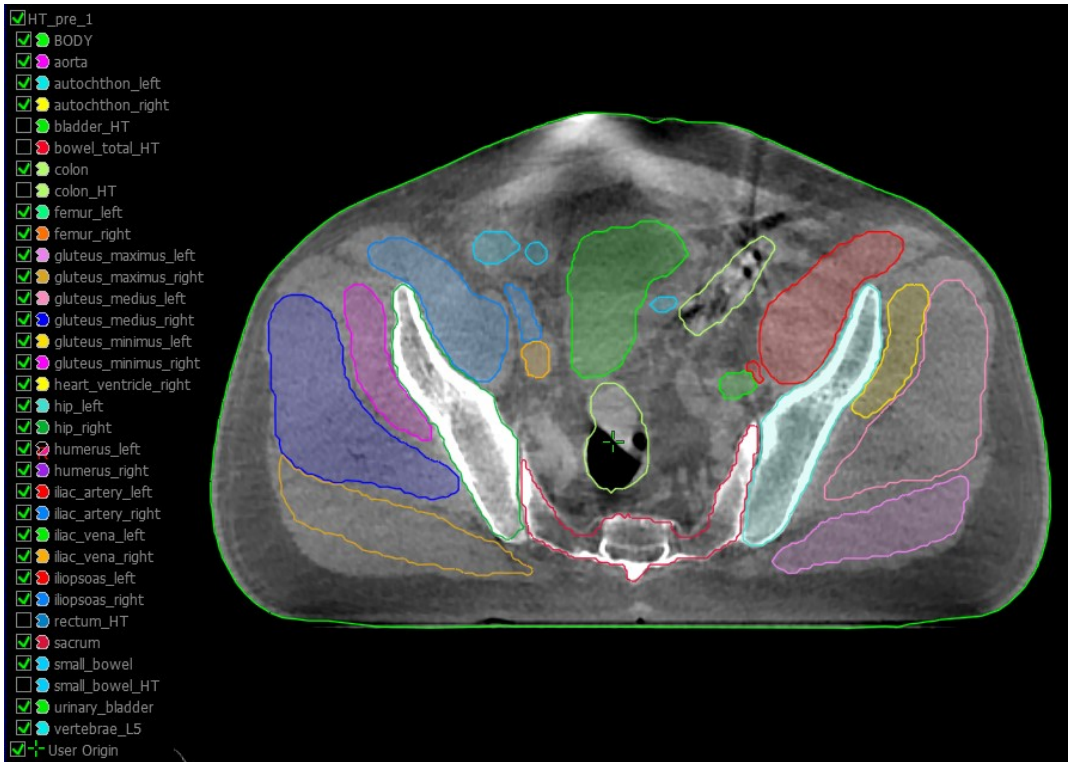

Figure 3.1: A CBCT slice in one patient including TotalSegmentator's segmentations checked off in the list. The remaining structures are copies revised by delineator one, not visualised on the scan.

TS distinguishes between small bowel and colon, but does not fully segment rectum. A bowel structure, consisting of small bowel and colon combined, was created as the union of TS's segmented small bowel and colon structures. It was observed that the algorithm frequently failed to include the entire colon or small bowel structure in its segmentation, particularly the structure walls, as illustrated for the colon in Figure 3.2. Therefore, an outer 3D margin of 2 mm was added to TS's bowel using the inbuilt margin function in Eclipse. Further, the bowel structure was post-processed in all 2D image planes. The structure boundaries were smoothed using the inbuilt smoothing function in Eclipse with smoothing level 2 to make the structure more realistic. Single structures smaller than 0.50 cm^2 were removed using the inbuilt clean up filter in Eclipse because the Python script used for geometrical comparisons was not able to perform calculations on such small structures. TS's small bowel, colon and bowel structures were adjusted to start inferiorly in the same slice as delineator one's most inferior segmentation and ended superiorly in the same slice as delineator one's most superior segmentation of the corresponding structure. When referring to the auto-segmentations by TS further in this thesis, it always implies inclusion of these post-processing steps.

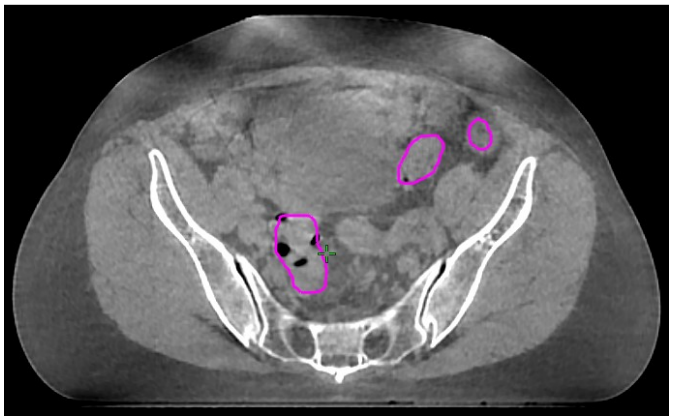

Figure 3.2: A CBCT slice in one patient with the colon segmented by TotalSegmentator. It is observed that the edges of the segmentations do not coincide with the colon walls. This is particularly evident for the colon loop in the upper right of the CBCT, i.e. the descending colon.

3.3 Geometric evaluation

For each patient, the total delineated bowel volume for each segmentation were extracted from Eclipse. To compare structures segmented on different CBCT scans, structures segmented on post-CBCTs were copied to the corresponding pre-CBCTs, and structures segmented on CBCTs of fx5 were copied to the corresponding CBCTs of fx1. Prior to all calculations, the pairwise compared bowel segmentations were adjusted to have an equal cranial extent due to varying CBCT scan lengths. In order to compare different volume segmentations and hence evaluate segmentation performances and to quantify intra- and interfractional bowel motion, DSCs and HD95 values were calculated between segmented structures using an in-house Python script developed by Helge Egil Seime Pettersen at HUH. To quantify segmentation performances, DSCs and HD95 values were calculated between delineator one's, delineator two's and TS's segmented structures of the bladder, small bowel, colon, bowel and rectum. All segmented structures were exported to a PC as DICOM RT Structure sets before the calculations were performed. To quantify intrafractional bowel motion, DSCs and HD95 values were calculated between manually bowel loop segmentations in the pre- and post-CBCTs in both fx1 and fx5. To quantify interfractional bowel motion, the same metrics were calculated between the manually bowel loop segmentations in the pre-CBCTs in fx1 and fx5. The script calculates DSCs and HD95 values between two structure segmentations slice by slice in all CT or CBCT slices where both segmentations are present and segmentations present in only one structure will be ignored.

The DSC, measuring the overlap between two 2D segmentations A and B, is defined by Equation 3.1, illustrated in Figure 3.3 [79]. In our study, the DSC in Equation 3.1

is calculated between two 2D segmentations, slice by slice for the transversal CBCT slices if both segmentations are present in the slice. |A| and |B| indicate areas and ∩ is the intersection operator.

$$\mathrm{DSC} = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{3.1}$$

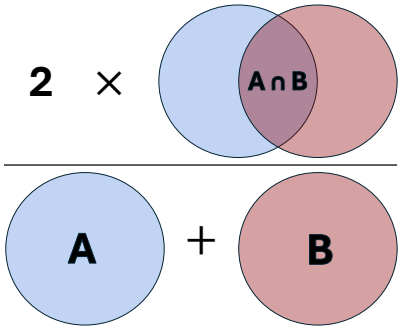

Figure 3.3: Illustration of the Dice Similarity Coefficient (DSC) calculated between two segmented circular structures A and B.

The HD95 value between two 2D finite point sets A and B, which in our study are segmented structures A and B, is defined by the 95$^{\text{th}}$ percentile of the distances $HD(A,B)$ determined by Equation 3.2. The distances are calculated between the outer borders of the structures A and B, as illustrated in Figure 3.4 [79]. As for the DSC, Equation 3.2 and 3.3 are calculated between the outer borders of two 2D segmentations defined within the same transversal CBCT slice.

$$HD(A,B) = \max(h(A,B), h(B,A)) \tag{3.2}$$

$h(A,B)$ is the directed Hausdorff Distance between point sets A and B, defined by equation 3.3 [79].

$$h(A,B) = \max_{a \in A} \min_{b \in B} ||a - b|| \tag{3.3}$$

$||a-b||$ is the distance between any point a in point set A and any point b in point set B. The distance can be measured as for instance Euclidean distance [79]. As expressed by Equation 3.3, the directed HD measures the maximum distance between point a in point set A to its closest point b in another point set B, as illustrated in Figure 3.4 where the shortest Euclidean distance is computed from all points on border A to all points on border B and vice versa [102]. Further, taking the maximum of the directed HDs as expressed in Equation 3.2 gives the HD between point sets A and B. In Figure 3.4, the 95$^{\text{th}}$ percentile of all the distances is returned as the HD95 value [102].

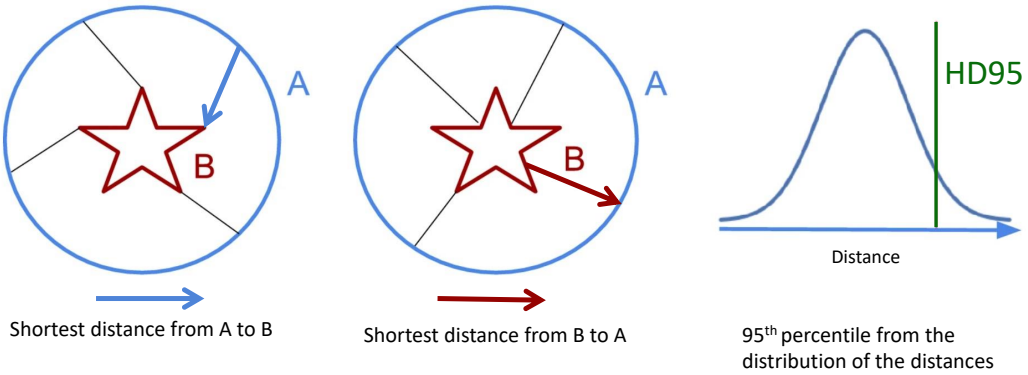

Figure 3.4: Illustration of how the 95th percentile Hausdorff Distance (HD95) is computed between two point sets A and B representing a segmentation of a circle and a star, respectively. Illustration from [102], edited.

3.4 Dosimetric evaluation

3.4.1 Impact of intrafractional bowel motion and inaccurate AI segmentations on estimated bowel exposure

Treatments were delivered as online adaptive SCRT in five fractions of 5.0 Gy each with a total dose of 25.0 Gy. CTVs were delineated around the RT targets including the primary tumour in the rectum, malignant tissue and diseased lymph nodes. The CTVs were expanded with an additional margin of 4 mm creating the PTVs. The treatment plans were simulated with 12 fields IMRT using 6 MV, which is the clinical practice of online ART of rectal cancer at HUH, with prescribed dose normalised to the median, meaning that 100% of the prescribed dose covered 50% of the target volume. The treatments were simulated as delivered with the Ethos system and evaluated in the "External Beam Planning" module in Eclipse. An example of a dose plan and a simulated dose distribution are illustrated in Figure E6.6.

The dose distribution was calculated on the synthetic CTs (sCT) using the Acuros (version 1.1.2.44) algorithm. For all patients, dose to the manual bowel segmentations on the pre- and post-CBCTs of both fx1 and fx5, and TS's bowel segmentations on the pre-CBCTs of fx1 were analysed. The DVH parameters of the absolute volumes receiving at least 5.0 Gy (V5), 7.5 Gy (V7.5), 8.9 Gy (V8.9), 10.0 Gy (V10), 12.6 Gy (V12.6), 20.0 Gy (V20) and 22.5 Gy (V22.5) were extracted. In each fraction, the volume difference (cc) between the manually segmented pre- and post-treatment bowels were calculated for all dose levels. In fx1, the volume difference between delineator one's and TS's segmented pre-treatment bowels were calculated for all dose levels.

3.4.2 Impact of bladder filling on bowel exposure

For each patient, the bladder volume in the pre- and post-CBCT of fx1 and fx5 were extracted from Eclipse. The absolute and relative intrafractional bladder volume difference in both fx1 and fx5 were calculated in addition to the absolute interfractional difference between fx1 and fx5. Any relation between the pre-treatment bladder volumes and the dose to the pre-treatment bowels was examined. It was studied if the absolute intrafractional change in bladder volume affected the difference between the pre- and post-treatment bowel V10 and V20 in both fx1 and fx5.

The correlations between the pre-treatment bladder volumes and the pre-treatment bowel V10s and V20s were determined using the Pearson correlation coefficient in fx1 and fx5 combined. Pearson correlation coefficients were also found between the intrafractional differences in pre- and post-treatment bladder volumes and the intrafractional differences in pre- and post-treatment bowel V10s and V20s, for both fractions. The Pearson correlation coefficients (r) were calculated using the "pearsonr" statistical function from the SciPy library in Python, measuring the linear relationship between the studied data. The coefficient ranges from -1 to +1, where $r = 0$ indicates no linear relationship between the bladder and irradiated bowel volume and $r = \pm 1$ indicates an exact linear relationship [103]. A hypothesis test of the null hypothesis of uncorrelated data were also performed using the same statistical function in Python.

Linear regression analyses were performed to relate the pre-treatment bladder volumes to the pre-treatment bowel V10s and V20s in both fractions. Similarly, the intrafractional differences between the pre- and post-treatment bladder volumes were related to the differences between the pre- and post-treatment bowel V10s and V20s. Linear regression analyses were performed using the scikit-learn library in Python. For each analysis, a linear model, described by Equation 3.4, was fitted to the data points, and the optimal coefficients β_0 and β_1 that best described the linear relationship between the studied data points were determined.

$$f(x) = \beta_0 + \beta_1 x \tag{3.4}$$

In Equation 3.4, f(x) represents bowel volume and x represents bladder volume. The coefficients β_0 and β_1 are the intercept and slope, respectively. β_0 represents the expected bowel volume or change in bowel volume when the bladder volume is empty or when there are no intrafractional change in the bladder volume, respectively. β_1 indicates how the bowel volume or change in bowel volume is influenced by the bladder volume or change in bladder volume, respectively [103].

3.5 Predicting grade ≥ 3 acute toxicity in bowel loops

How bowel motion influence severe bowel toxicity estimates was analysed using dose-response modelling. Banerjee et al. analysed the relation between dose exposure and small bowel toxicity of grade ≥ 3 in rectal cancer patients treated with long course neoadjuvant CRT [92]. They found that every 5 Gy-interval between 5 and 40 Gy were significantly associated with toxicity, regardless of defining the bowel as individual loops or the peritoneal space [92]. Logistic regression curves for toxicity probability estimation were calculated for 15 Gy and 25 Gy, shown in Figure 3.5 for small bowel loop delineation for the lowest dose level [92].

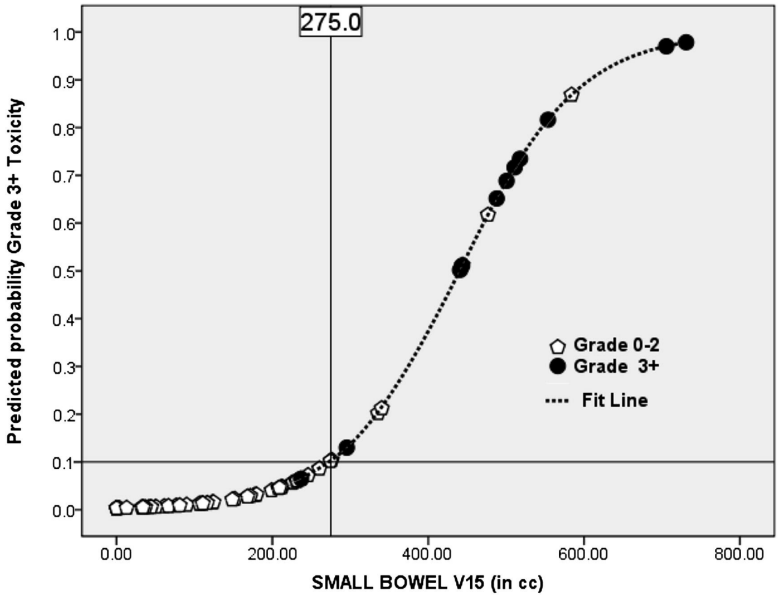

Figure 3.5: Logistic regression curve for toxicity probability estimation for small bowel volume receiving at least 15 Gy (V15) from Banerjee et al. [92]. The y-axis represents predicted probability of grade ≥ 3 acute small bowel toxicity. Markers indicate the grade of toxicity for each patient in their study. The crossed lines illustrate that small bowel V15 < 275 cc was associated with a $< 10\%$ risk of grade ≥ 3 acute small bowel toxicity [92].

The general equation of a logistic regression curve is the logit function in Equation 3.5, where $p(x)$ is the probability related to the value of x by the logit function [103]. The coefficients β_0 and β_1 are the intercept and the slope respectively of the log-odds in Equation 3.6 where $\frac{p(x)}{1-p(x)}$ is the odds of the logit function [103]. Using WebPlotDigitizer [104], numerical data from Figure 3.5 were extracted. These data points paired each small bowel V15 with the predicted toxicity probability, and were imported into Python to recreate the logistic regression curve using the statsmodels library. β_0 and β_1 were found and the equation of the logistic regression curve in Figure 3.5 was determined.

$$p(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{3.5}$$

$$\ln \frac{p(x)}{1-p(x)} = \beta_0 + \beta_1 x \tag{3.6}$$

The EQD_2 formula (Equation 2.9) was first used to convert a dose of 15 Gy delivered with a dose per fraction of 1.8 Gy to the isoeffective dose in 2 Gy fractions. Then, the isoeffective dose in 2 Gy fractions was converted to the equivalent dose in 5 Gy fractions. The generic α/β value of 10 Gy for acute toxicity was used in both calculations. The calculations resulted in an equivalent dose of 12.6 Gy using 5 Gy fractions.

Holyoake et al. performed a meta-analysis of publications reporting small bowel dose-volumes and toxicity data in rectal cancer patients treated with long course neoadjuvant CRT, including the publication by Banerjee et al. [4]. They found that each of the dose-volume measures between small bowel V5 and V40 in 5 Gy-intervals were statistically significantly associated with grade ≥ 3 acute small bowel toxicity [4]. Logistic regression curves for toxicity probability estimation were calculated for each small bowel dose-volume parameter in 5-Gy intervals from V5 to V50. The given logistic regression parameters for V10 were $\beta_0 = -2.63$ and $\beta_1 = 0.005$ [4], which were used to recreate the logistic regression curve in Python. The publications studied by Holyoake et al. differed in the number of delivered RT fractions and the total RT dose [4]. To find the equivalent dose in 5 Gy fractions using SCRT, we assumed that the V10 provided by Holyoake et al. was based on RT delivery of 1.8 Gy/fraction delivered in 28 fractions. Using the EQD_2 formula with the generic α/β value of 10 Gy for acute toxicity, a dose of 10 Gy delivered as 28×1.8 Gy is equivalent to 8.9 Gy delivered as 5×5 Gy.

The logistic regression equations were utilised to predict the probability ($p(x)$) of grade ≥ 3 acute toxicity in the pre- and post-treatment bowel loops receiving at least 12.6 Gy (x) using the model by Banerjee et al. and 8.9 Gy using the model by Holyoake et al. for each patient in fx1 and fx5. Additionally, the intra- and interfractional differences in probabilities of grade ≥ 3 acute bowel toxicity were examined for all patients. To quantify intrafractional differences in predicted toxicity probabilities, the difference in predicted toxicity probability for the pre- and post-treatment bowels were calculated in both fx1 and fx5 for all patients. Interfractional differences in predicted toxicity probabilities were investigated by calculating the difference between the predicted toxicity probability for the pre/post-treatment bowels in fx1 and the pre/post-treatment bowels in fx5 for all patients.

3.6 Hypothesis testing

Statistical analyses were performed for hypothesis testing and plots were generated to justify the choices of statistical tests in R (version 4.2.2) and Python (version 3.11.5). A one- or two-sided (depending on the alternative hypothesis) p-value of $p < 0.05$ was considered statistically significant for all statistical tests. If a test resulted in a p-value ≤ 0.05, the null hypothesis (H_0) was rejected and there was a statistically significant difference between the studied samples. If $p > 0.05$, the alternative hypothesis (H_a) was rejected meaning that the test failed to show any difference between the two data sets.

The Mann-Whitney U test was performed to investigate statistically significant differences in segmentator performances and intrafractional bowel motion between fx1 and fx5, furthermore to compare intra- and interfractional bowel motion. Since the segmentations, performed by different delineators, are segmentations of the same structures on the same CBCT images, the data suggested a paired statistical test such as the paired Wilcoxon signed rank test. However, a paired statistical test was not used because of unequal sample sizes due to different scan lengths. The Mann-Whitney U test is a test of the null hypothesis that the distributions underlying the two independent samples it compares differ by a location shift of Δ_0. In our tests, $\Delta_0 = 0$. H_0 is tested against an alternative hypothesis that the distributions differ by any other location shift [103]. The tests were performed as stated below for DSCs and HD95 values with μ representing mean values.

Manual segmentations vs. TS or intrafractional bowel motion

$$H_{0_{\text{DSC}}} : \mu_{\text{DSC manual}} - \mu_{\text{DSC TS/intra}} = 0$$
$$H_{a_{\text{DSC}}} : \mu_{\text{DSC manual}} - \mu_{\text{DSC TS/intra}} > 0$$

$$H_{0_{\text{HD95}}} : \mu_{\text{HD95 manual}} - \mu_{\text{HD95 TS/intra}} = 0$$
$$H_{a_{\text{HD95}}} : \mu_{\text{HD95 manual}} - \mu_{\text{HD95 TS/intra}} < 0$$

TS vs. intrafractional bowel motion

$$H_{0_{\text{DSC}}} : \mu_{\text{DSC TS}} - \mu_{\text{DSC intra}} = 0$$
$$H_{a_{\text{DSC}}} : \mu_{\text{DSC TS}} - \mu_{\text{DSC intra}} < 0$$

$$H_{0_{\text{HD95}}} : \mu_{\text{HD95 TS}} - \mu_{\text{HD95 intra}} = 0$$
$$H_{a_{\text{HD95}}} : \mu_{\text{HD95 TS}} - \mu_{\text{HD95 intra}} > 0$$

Intra- vs. interfractional bowel motion

$$H_{0_{\text{DICE}}} : \mu_{\text{DICE intra}} - \mu_{\text{DICE inter}} = 0$$
$$H_{a_{\text{DICE}}} : \mu_{\text{DICE intra}} - \mu_{\text{DICE inter}} > 0$$

$$H_{0_{\text{HD95}}} : \mu_{\text{HD95 intra}} - \mu_{\text{HD95 inter}} = 0$$
$$H_{a_{\text{HD95}}} : \mu_{\text{HD95 intra}} - \mu_{\text{HD95 inter}} < 0$$

Intrafractional bowel motion fx1 vs. fx5

$$H_{0_{\text{DICE}}} : \mu_{\text{DICE fx1}} - \mu_{\text{DICE fx5}} = 0$$
$$H_{a_{\text{DICE}}} : \mu_{\text{DICE fx1}} - \mu_{\text{DICE fx5}} \neq 0$$

$$H_{0_{\text{HD95}}} : \mu_{\text{HD95 fx1}} - \mu_{\text{HD95 fx5}} = 0$$
$$H_{a_{\text{HD95}}} : \mu_{\text{HD95 fx1}} - \mu_{\text{HD95 fx5}} \neq 0$$

The paired Wilcoxon signed rank test and Levene's test were performed to assess whether the volume differences between the manually segmented pre- and post-treatment bowels receiving at least 5.0 Gy, 7.5 Gy, 20.0 Gy and 22.5 Gy differed in distribution or variance between fx1 and fx5. The dose levels 5.0 Gy and 7.5 Gy were chosen to represent low doses, and 20.0 Gy and 22.5 Gy to represent high doses. A paired test was used since the volume differences in fx1 and fx5 are found for the same 15 patients, separated by time. The paired Wilcoxon signed rank test was also performed to investigate any statistically significant difference in dose delivered to delineator one's and TS's segmented pre-treatment bowels in fx1 for the same chosen dose levels. Also, it was used to test for statistical significant differences between the different segmentators' total segmented bowel volumes in the post-CBCTs of fx1. In the latter two comparisons, a paired test was used because the same structures were segmented in the same CBCTs by different delineators. Additionally, it was performed to test for statistically significant differences in the predicted probabilities for grade ≥ 3 acute bowel toxicity between the models provided by Banerjee et al. and Holyoake et al. Here, a paired test is suitable because the toxicities are predicted for the same bowel loops using different models. When investigating the intra- and interfractional differences in toxicity estimates, the test was performed on the absolute values of the differences. The paired Wilcoxon signed rank test investigates paired data $(x_1, y_1), ..., (x_n, y_n)$ where x_m and y_m, $m \in [1, n]$, can for instance be data points from fx1 and fx5. It tests hypotheses about the mean value μ_D of the differences $D_m = x_m - y_m$ with a null hypothesis that the D_m's are symmetric about some value Δ_0, which in these cases equals zero [103]. H_0 is tested against H_a that the differences are symmetric about some other value, not equal to Δ_0 [103]. All paired Wilcoxon signed rank tests were performed as stated below.

DVH parameters for manual segmented bowels from fx1 vs. fx5

$$H_{0_{\text{V5/V7.5/V20/V22.5}}} : \mu_D = 0$$
$$H_{a_{\text{V5/V7.5/V20/V22.5}}} : \mu_D \neq 0$$

DVH parameters for delineator one's vs. TS's bowel segmentations

$$H_{0_{\text{V5/V7.5}}} : \mu_D = 0$$
$$H_{a_{\text{V5/V7.5}}} : \mu_D > 0$$

$$H_{0_{\text{V20/V22.5}}} : \mu_D = 0$$
$$H_{a_{\text{V20/V22.5}}} : \mu_D < 0$$

Total segmented bowel volume by delineator one vs. delineator two

$$H_0 : \mu_D = 0$$
$$H_a : \mu_D \neq 0$$

Total segmented bowel volume by delineator one vs. TS

$$H_0 : \mu_D = 0$$
$$H_a : \mu_D > 0$$

Predicted toxicities using the model provided by Banerjee et al. vs. Holyoake et al.

$$H_{0_{\text{(pre fx1/post fx1/pre fx5/post fx5/intra fx1/intra fx5/inter pre/post fx1-fx5)}}} : \mu_D = 0$$
$$H_{a_{\text{(pre fx1/post fx1/pre fx5/post fx5/intra fx1/intra fx5/inter pre/post fx1-fx5)}}} : \mu_D < 0$$

According to the studies of Heijkoop et al. [105] and Ahmad et al. [106], a bladder volume of approximately 300 ml which equals 300 cc are assumed to correspond with a full bladder. The two-sample t-test and the F-test were utilised to investigate if the patients with an approximately full bladder had different bowel V10 and V20 than the patients with a smaller bladder filling. Only the pre-treatment bladders and bowels of fx1 and fx5 were studied, based on Figure 4.25. The t-test analyses the difference in means of two independent samples, and the F-test compares their variances (σ). The two-sample t-test assumes equal variances which can be assessed using the F-test [103]. A null hypothesis that the samples have identical means (t-test) and variances (F-test) are tested against an alternative hypothesis that the samples differ in mean and variance respectively.

Bowel V10 and V20 for full vs. small bladders

$$H_{0_{\text{V10}}} : \mu_{(\text{bladder}\geq 300\text{ cc})} - \mu_{(\text{bladder}<300\text{ cc})} = 0$$
$$H_{a_{\text{V10}}} : \mu_{(\text{bladder}\geq 300\text{ cc})} - \mu_{(\text{bladder}<300\text{ cc})} < 0$$

$$H_{0_{\text{V20}}} : \mu_{(\text{bladder}\geq 300\text{ cc})} - \mu_{(\text{bladder}<300\text{ cc})} = 0$$
$$H_{a_{\text{V20}}} : \mu_{(\text{bladder}\geq 300\text{ cc})} - \mu_{(\text{bladder}<300\text{ cc})} > 0$$

$$H_0 : \sigma^2_{(\text{bladder}\geq 300\text{ cc})} - \sigma^2_{(\text{bladder}<300\text{ cc})} = 0$$
$$H_a : \sigma^2_{(\text{bladder}\geq 300\text{ cc})} - \sigma^2_{(\text{bladder}<300\text{ cc})} \neq 0$$

It is uncertain whether the bowel motion and thus potentially the volume difference in fx5 is dependent of the motion in fx1 and vice versa. De Jong et al. argued that larger interfractional variations of the mesorectum, which is part of the target volume, and OARs such as the bowel bag cause data of different fractions within one patient to be considered independent [75]. Whether the bowel motion in fx1 and fx5 are dependent can affect the Mann-Whitney U test and Levene's test, which require independent data [103]. Individual value plots, one illustrated in Figure A6.3, were created to assess whether compared samples seemed to be dependent or not.

Q-Q Plots and the Shapiro-Wilk test were utilised to test for normality as required by the t-test and F-test [103]. Example of Q-Q plots are shown in Figure A6.2 where sample quantiles from the volume differences for a studied dose and fraction are plotted against theoretical quantiles from the normal distribution, forming a scatter plot. A line passes through the first quartile (Q1) and the third quartile (Q3) of the normal distribution [107]. If the points of the scatter plot aligns with the line, the data can be considered normally distributed. The Shapiro-Wilk test is a common statistical test used to examine whether data are normally distributed, appropriate for small datasets with less than 50 samples [107]. The Mann-Whitney U test, the paired Wilcoxon signed rank test and/or Levene's test were used in hypothesis testing for data that failed to show normality.

Density plots were created to visually inspect the distribution and symmetry of the data. The Mann-Whitney U test assumes that the underlying distributions of the compared samples have the same shape and spread with a possible difference in mean values. The paired Wilcoxon signed rank test assumes continuity and symmetry of the compared data [103]. Figure A6.1 illustrates a density plot together with a line indicating the mean value. If a distribution is symmetric, the mean is located at the center of the distribution with approximately half of the data distributed with similar shape on each side of the mean.

Chapter 4

Results

4.1 Segmentation performance

4.1.1 Segmentations

The total segmented bowel volume for all patients in the post-CBCTs of fx1 was similar for the manual segmentations (average 574.8 cc and 583.5 cc by delineator one and two, respectively) and lower for TS's segmentations (average 411.0 cc). An overview of the total segmented post-treatment bowel volumes of fx1 per patient, per segmentator, is given in Table B6.1. The paired Wilcoxon signed rank test showed no statistically significant differences between delineator one's and two's total segmented bowel volumes. However, the total segmented bowel volume by delineator one was statistically significantly greater than the volume segmented by TS ($p < 0.01$). These results imply that the two manual delineators identified similar bowel amounts in the CBCTs, while TS failed to identify greater amounts of the bowels. The total segmented bowel volumes were found for the post-CBCTs of fx1 since these are the only CBCTs segmented by all three segmentators.

The main challenges experienced when manually segmenting the structures of interest were caused by the low image quality of the CBCTs. As mentioned in Chapter 2, CBCTs are known to have reduced image quality compared to conventional CTs. Several of the CBCTs included in our study were to a greater or lesser extent occupied by artifacts, often due to air in the bowels. Observations revealed that almost all studied CBCTs had increased amounts of artifacts in the most cranial slices. This was caused by multiple patients frequently having greater amounts of air in their bowels located in the cranial region of the pelvis compared to the caudally located bowels. Therefore, identifying the cranial bowel loops was challenging, and the manual segmentators segmented the bowel in these slices resembling a bowel bag structure, including possible

loops hidden by the artifacts. In fx1, the trend of increasing amounts of artifacts in the most cranial slices had less influence on the image quality and thus the segmentation performances for patients ER1, ER8, ER11 and ER12. This is reflected in the segmentation performances' plots (Figure 4.5 and 4.6) where, for the four mentioned patients, all median DSCs are ≥ 0.80 with small interquartile ranges (IQRs). Also, their HD95 values are relatively low compared to the corresponding results for the other patients except for ER8. Figure 4.1 illustrates a cranial and a caudal CBCT slice from the pre-CBCT scan of patient ER14 in the first fraction, with and without the bowels segmented by delineator one. In the cranial slice (Figure 4.1 a), the bowel contains multiple air bubbles resulting in streak artifacts. To ensure that all bowel loops were included in the segmentation, delineator one and two segmented the bowel in this slice as a bowel-bag-like structure (Figure 4.1 b). The image is clearer in the caudal CBCT slice (Figure 4.1 c), consequently, the bowel loops are easy to identify and segment (Figure 4.1 d). The challenges in segmenting the cranial loops were experienced by both delineator one and two. Mostly, it was not experienced that either of the manual delineators managed to identify more loops than the other. TS, however, did not expand its bowel segmentations in the presence of artifacts. For the cranial post-CBCT slice of ER14 in Figure 4.1 a), the manually bowel segmentations were almost equal resulting in a DSC of 0.91 and HD95 value of 7.70 mm for the current slice. TS's bowel segmentation, however, diverged from delineator one's, resulting in a DSC of 0.55 and HD95 value 24.42 mm, indicating poorer segmentation performance by TS than by the manual segmentators in the current slice. Some patients had CBCT scans more occupied by artifacts than others. Examples are ER3 and ER13, illustrated in Figure 4.2, where artifacts were present in the caudal as well as the cranial slices. In these patients, delineator one and two did not attempt to distinguish between the small bowel and colon in fx1, and segmented the bowel in a bowel-bag-like structure (see Section 3.2.2). However, TS attempted to segment the small bowel and colon loops, and the artifacts caused TS to exclude large parts of the bowels. These segmentation uncertainties influenced the DSCs and HD95 values for ER3 and ER13 similar to the example of ER14, resulting in the poorer results of the metrics comparing the segmentations by delineator one and TS than the results comparing the manual segmentations. The segmentation performances' plots (Figure 4.5 and 4.6) reflect such segmentation disagreements where, the comparison between delineator one's and TS's bowel segmentations resulted in lower median DSCs and higher median HD95 values than between the manual bowel segmentations for the majority of the patients.

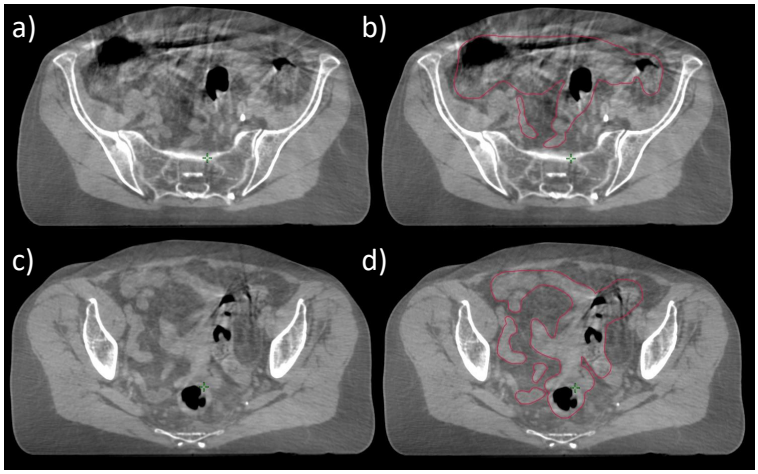

Figure 4.1: Cranial and caudal pre-CBCT slices of patient ER14 in fraction two illustrating differences in artifacts. a) CBCT slice with poor image quality due to artifacts, cranially located within the pelvis. b) Bowel from a) segmented by delineator one. Due to artifacts, identifying the bowel loops was challenging. Hence, the bowel is segmented in a bowel-bag-like structure. c) CBCT slice located more caudally within the patient, less occupied by artifacts. The bowel loops were easily identified and defined. d) Bowel loops from c) segmented by delineator one.

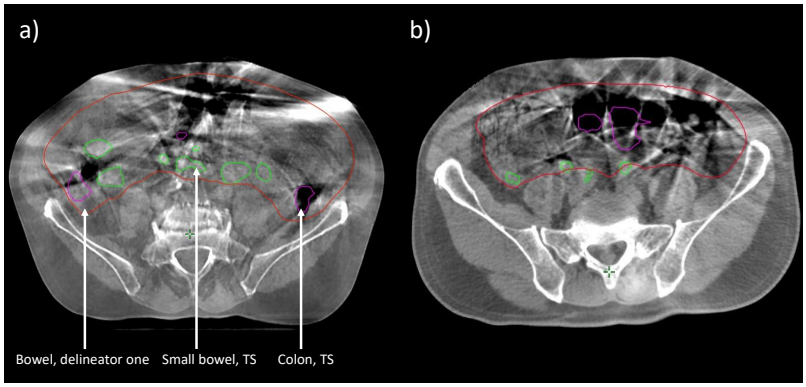

Figure 4.2: A CBCT slice of patient a) ER13 and b) ER3 occupied by artifacts. The bowel (red) is segmented by delineator one and the small bowel (green) and colon (magenta) by TotalSegmentator (TS).

During manual segmentation, both delineators experienced challenges in distinguishing between the small bowel, colon and rectum, which were worsened by the poor image quality. Delineator one observed that TS also struggled to distinguish between the small bowel and colon, with poorer segmentation performances for bowels containing air. TS especially struggled to identify the ascending colon as colon, often mislabelling it as small bowel. The majority of the patients mostly had a bowel anatomy that initially resembled the "theoretical" anatomy illustrated in Figure 2.25. However, both the small bowel and colon loops in individual patients could bend in various ways, making it even more challenging to distinguish between the different bowel structures. For this reason, delineator one and two in some patients interpreted

the bowel anatomy differently, which could affect the DSCs and HD95 values between their segmentations. Figure 4.3 shows an example of a post-CBCT slice of patient ER7 in fx1 where delineator one and two disagreed in the distinction between the small bowel and colon. In Figure 4.3 a), the output from the Python script calculating DSCs and HD95 values illustrates where the two delineators segmented colon, where their segmentations overlap, and where they differ. The yellow area indicates where delineator one defined the visible bowel structure as colon while delineator two defined the bowel segment as small bowel. The orange area, which almost not occurs in this slice, represents where only delineator two segmented colon. Figure 4.3 b) shows a plot of HD95 values against DSCs calculated between delineator one's and two's segmentations for the small bowel, colon and bowel in the slice illustrated in Figure 4.3 a). The optimal location of each HD95 vs. DSC point is in the lower right corner of the plot, where the DSC is close to one and the HD95 value close to zero, indicating a high degree of agreement between the segmentations. The illustrated small bowel and colon segmentations demonstrate poorer DSC and HD95 results than the median values between the manual segmentations for the corresponding structures in all patients (Table 4.1). However, when the small bowel and colon segmentations were combined to a joint bowel structure, the geometric evaluation metrics were improved, as illustrated in Figure 4.3 b).

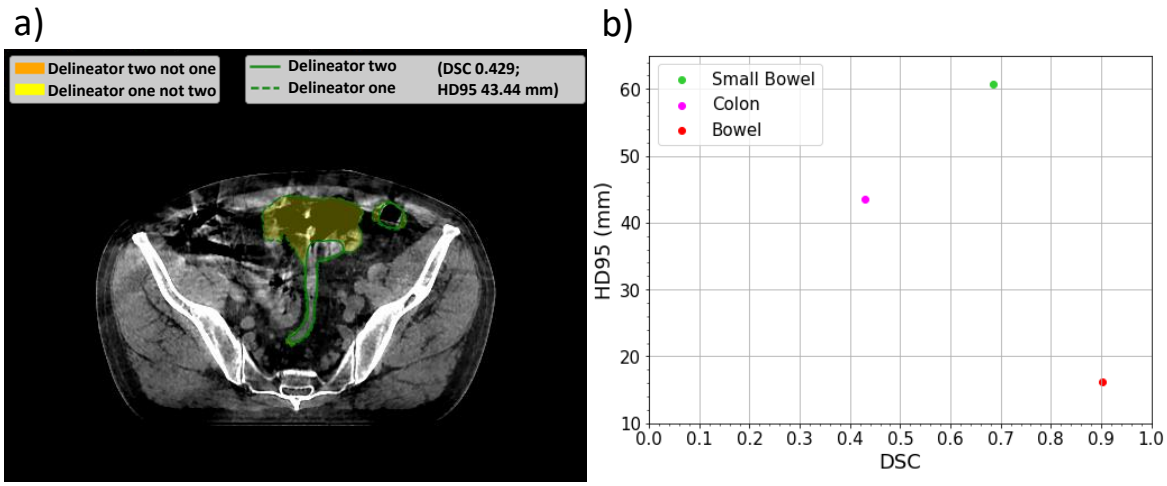

Figure 4.3: a) Output from the Python script calculating Dice Similarity Coefficients (DSCs) and 95th Hausdorff Distances (HD95s) between delineator one's and two's colon segmentation in one CBCT slice of patient ER7. Coloured are represents non-overlapping area. b) Scatter plot of HD95 values against DSCs for the small bowel (green), colon (magenta) and bowel (red) in the CBCT slice in a), calculated between segmentations performed by delineator one and two.

Some patients had greater amounts of air in their rectum, making it challenging to identify the transition between the colon and the rectum. This could have affected both

the segmentation evaluation and the measured intra- and interfractional bowel motion if the colon's inferior extent varied due to segmentation uncertainties and not actually motion. An example from patient ER6 is shown in Figure 4.4. However, since the Python script calculating the 2D DSCs and HD95 values only performed calculations if both segmentations of a structure were present in the slice, this did not affect the metrics.

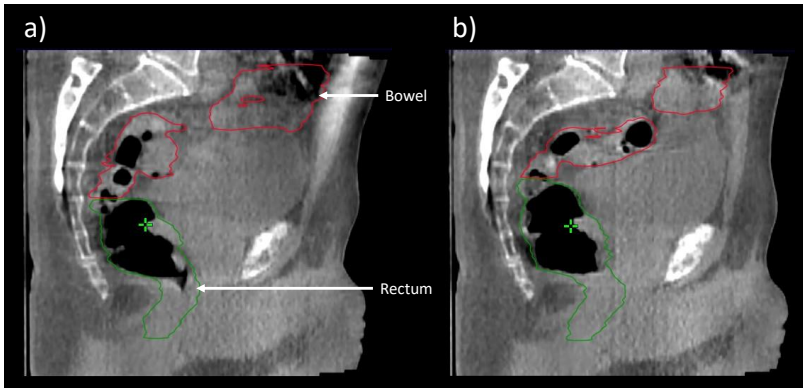

Figure 4.4: Pre-CBCT slice (a) and corresponding post-CBCT slice (b) of patient ER6 illustrating the bowel (red) and rectum (green) containing air, making it challenging to detect the transition between them.

4.1.2 Geometric evaluation of segmentation performance

The geometric metrics used in evaluating segmentation performances indicated that the manual delineators were more able to identify the bladder, the small bowel loops and the colon loops in CBCTs than TS. Table 4.1 shows the median (first quartile (Q1):third quartile (Q3)) DSCs and HD95 values for the bladder, small bowel, colon and bowel for the different segmentation comparisons of the structures in the pre- and post-CBCTs from the first fraction in all 15 patients. For all structures, the DSCs between delineator one's and two's segmentations were consistently higher than those between delineator one's and TS's segmentations while the HD95 values were consistently lower.

Table 4.1: Median Dice Similarity Coefficients (DSCs) and 95th percentile Hausdorff Distances (HD95s) measuring segmentation performances. TS refers to TotalSegmentator. Q1 and Q3 refers to the first and third quartile, respectively. Pre and post refers to segmentations of the structures on the pre- and post-CBCTs, respectively, in fraction one.

Segmentation		**DSC**	**HD95**
Structure	**Evaluation**	**Median (Q1:Q3)**	**Median (Q1:Q3) (mm)**
Bladder	Delineator one vs. two, post	1.00 (1.00:1.00)	0.00 (0.00:0.00)
	Delineator one vs. TS, pre	0.94 (0.90:0.97)	4.46 (3.04:7.74)
	Delineator one vs. TS, post	0.95 (0.87:0.97)	4.08 (2.88:8.01)
Small Bowel	Delineator one vs. two, post	0.85 (0.74:0.89)	11.89 (6.77:19.37)
	Delineator one vs. TS, pre	0.65 (0.41:0.77)	19.96 (12.01:31.97)
	Delineator one vs. TS, post	0.66 (0.40:0.80)	18.30 (10.52:34.26)
Colon	Delineator one vs. two, post	0.87 (0.73:0.92)	6.16 (3.04:37.72)
	Delineator one vs. TS, pre	0.76 (0.66:0.85)	11.65 (5.66:28.47)
	Delineator one vs. TS, post	0.77 (0.64:0.86)	13.47 (4.81:39.06)
Bowel	Delineator one vs. two, post	0.90 (0.87:0.93)	7.69 (4.81:12.83)
	Delineator one vs. TS, pre	0.80 (0.59:0.86)	14.42 (7.69:33.37)
	Delineator one vs. TS, post	0.83 (0.68:0.88)	12.50 (6.31:26.13)

For all segmentations, the bladder resulted in the highest DSCs and lowest HD95 values, implying that the bladder is easier to identify in CBCTs compared to the small bowel, colon and rectum. The rectum was only segmented by delineator one and two, resulting in a median (Q1:Q3) DSC of 0.94 (0.89:0.96) and HD95 value of 2.03 mm (1.36:3.27mm). The small bowel resulted in the lowest median DSCs and highest HD95 values, indicating that of all the investigated structures, the small bowel is most challenging to identify and segment in CBCTs. The colon resulted in higher median DSCs and lower HD95 values than the small bowel, for all segmentations. Further, the median DSCs for each segmentation of the bowel (small bowel and colon combined) were greater than those for the small bowel and colon. The median HD95 value for the bowel per segmentation comparison was always lower than the corresponding value for the small bowel, and in 1/3 of the cases lower than the value for colon. However, our study has uncertainties due to few patients, and differences in DSCs and HD95 values may not represent significant differences. Since the bowel generally resulted in better scores of the geometrical metrics, and due to the time spent on delineating the small bowel and colon separately in fx1, it was chosen to segment only the combined bowel in fx5.

The Mann-Whitney U test indicated statistically significantly higher DSCs between the manually segmented bowels by delineator one and two than both between delineator one's and TS's bowel segmentations and the DSCs measuring intrafractional bowel motion ($p < 0.01$) for the DSCs illustrated in Figure 4.5. The test also showed that the HD95 values between the manually segmented bowel loops were statistically significantly lower than both between delineator one's and TS's bowel segmentations and the

HD95 values measuring intrafractional bowel motion ($p < 0.01$) for the HD95 values illustrated in Figure 4.6. The DSCs between delineator one's and TS's bowel segmentations were statistically significantly lower than the DSCs measuring intrafractional bowel motion ($p < 0.05$). However, the HD95 values calculated between delineator one's and TS's bowel segmentations were not statistically significantly greater than those for intrafractional bowel motion. These results indicate that the manual delineation uncertainties are less than the difference observed between pre- and post-CBCTs, hence the bowel motion can be extracted. On the other hand, using TS unables these metrics to distinguish motion from delineation uncertainty in the studied CBCTs.

Figure 4.5 and 4.6 illustrate box plots of DSCs and HD95 values, respectively, for the bowel comparing segmentation performances in post-CBCTs and intrafractional bowel motion in each individual patient in fx1. In the figures, the highest DSCs and lowest HD95 values are found between the two manually delineators' bowel loop segmentations. In contrast, the lowest DSCs and highest HD95 values are mostly found between delineator one's and TS's segmentations. For each patient except ER3, the median DSCs calculated between the manual segmentations were higher than the corresponding medians measuring intrafractional bowel motion. The median HD95 values calculated between the manual segmentations were lower than the corresponding medians measuring intrafractional bowel motion in all patients except ER3, ER8 and ER9.

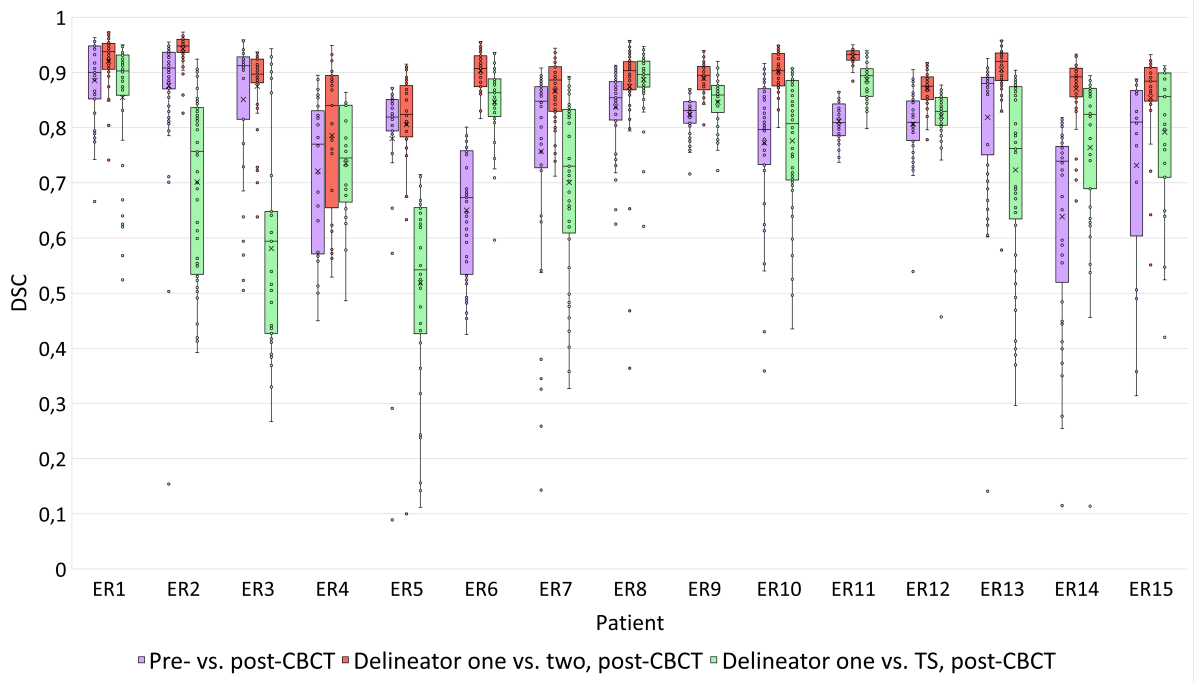

Figure 4.5: Segmentation performances plot; box plots of Dice Similarity Coefficients (DSCs) for bowel segmentations of fraction one in each individual patient, comparing intrafractional bowel motion and segmentation performances. Intrafractional bowel motion is illustrated in purple, where "Pre- vs. post-CBCT" refers to DSCs calculated between delineator one's bowel loop segmentations on each pre- and post-CBCT. Segmentation performances are measured between delineator one's and two's bowel loop segmentations on each post-CBCT (red) and between delineator one's and TS's bowel loop segmentations on each post-CBCT (green). Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Whiskers are ranging from the minimum to the maximum value that is not an outlier. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3.

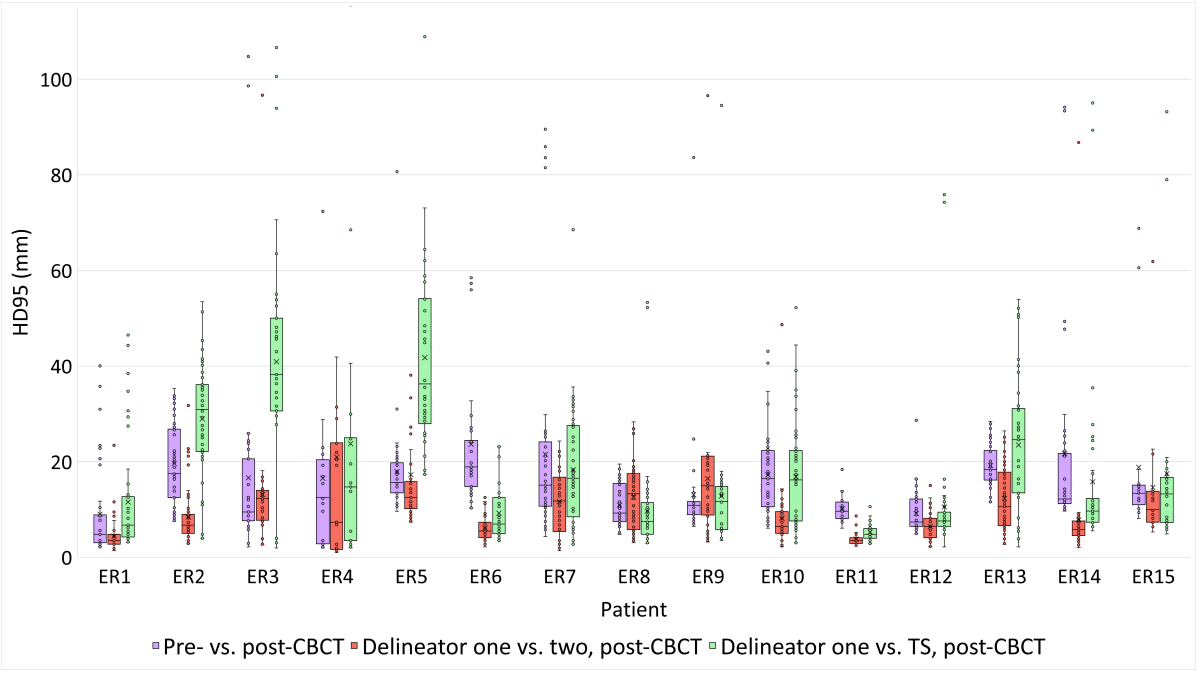

Figure 4.6: Segmentation performances plot; box plots of 95th Hausdorff Distances (HD95s) for bowel segmentations of fraction one in each individual patient, comparing intrafractional bowel motion and segmentation performances. Intrafractional bowel motion is illustrated in purple, where "Pre- vs. post-CBCT" refers to HD95s calculated between delineator one's bowel loop segmentations on each pre- and post-CBCT. Segmentation performances are measured between delineator one's and two's bowel loop segmentations on each post-CBCT (red) and between delineator one's and TS's bowel loop segmentations on each post-CBCT (green). Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Whiskers are ranging from the minimum to the maximum value that is not an outlier. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3.

The density plots in Figure A6.1 a) and b) illustrate an approximately similar shape for the distributions of the DSCs and HD95 values, respectively, for the bowel segmentation comparisons for all patients. The distributions of the DSCs appear to be negatively skewed while the distribution of the HD95 values appear to be positively skewed, consistent with the observations in Figure 4.5 and 4.6. All distributions have a slight difference in spread and their mean values are shifted relative to each other. The statistically significant differences between the DSCs for the manual bowel segmentations, the DSCs calculated between the bowel segmentations by delineator one and TS and those measuring intrafractional bowel motion are in good agreement with the observations from the density plots in Figure A6.1 a), where the peak of the distribution and the mean value for the manual segmentations are located at higher DSCs than the other illustrated peaks and mean values. The same applies to the HD95 values, where the statistically significant differences align with the observations from the density plots in Figure A6.1 b), where the peak of the distribution and the mean value for the manual segmentations are lower than the other illustrated peaks and mean values. The mean HD95 value between delineator one and TS is the highest illustrated mean value, however, it is located close to the mean value measuring intrafractional bowel motion, which may explain the absence of a statistically significant difference between their HD95 values.

In Figure 4.5 and 4.6, outliers are present in the majority of the patients' DSCs and HD95 values, both regarding segmentation performances and intrafractional bowel motion. While outliers from the intrafractional bowel motion results may be due to actual motion and thus true location displacement of the bowel, outliers from the segmentation comparisons represent segmentation uncertainties, caused by poor image quality and artifacts in the CBCTs. In some CBCT scans, the segmentators seemed to disagree on the cranial and caudal extent of the studied bowel and perhaps how much of a structure must be visible in order to segment it. Segmentation uncertainties in the most caudal segmented slices, in addition to the segmentation uncertainties in the cranial slices, often resulted in the outliers in the segmentation performances' plots. The CBCT slice in Figure 4.7 a) illustrates a discrepancy between delineator one and TS when segmenting the most inferior bowel loop when only a small part of the bowel was visible. The dose distribution within the same slice is illustrated in Figure 4.7 b).

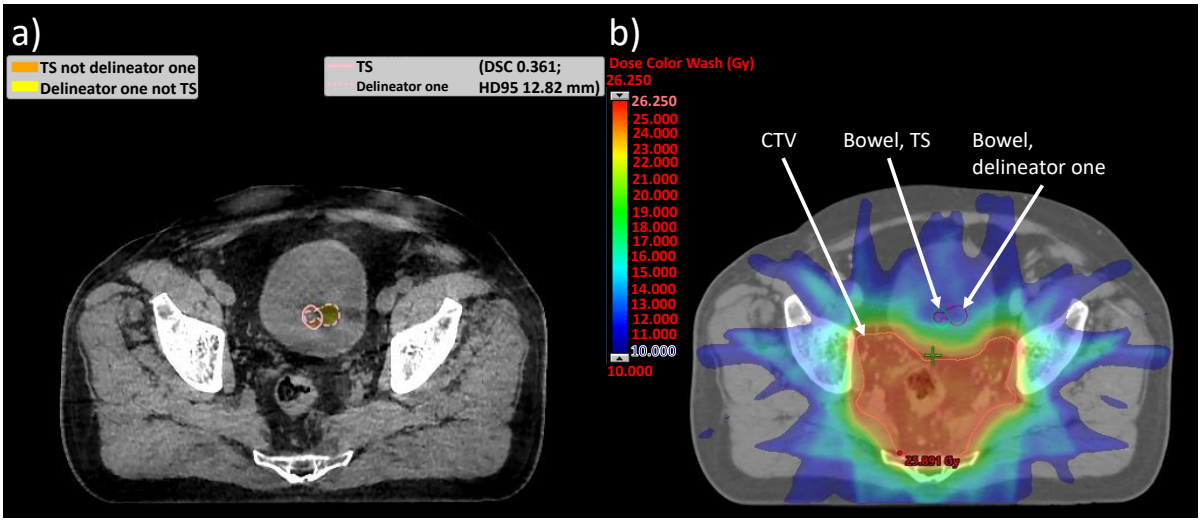

Figure 4.7: a) Output from the Python script calculating Dice Similarity Coefficients (DSCs) and 95th Hausdorff Distances (HD95s) between the bowel segmented by delineator one (pink dashed lines) and TotalSegmentator (AI) (pink solid lines) in one CBCT slice resulting in a low DSC. Coloured area represents non-overlapping area. b) Dose distribution within the same CBCT slice with delineated CTV (pink) and the bowel segmented by delineator one (red) and TotalSegmentator (AI, green).

The segmentation performances' plots (Figure 4.5 and 4.6) indicate variations in TS's segmentation performance in different patients. Figure 4.8 shows an example of a post-CBCT slice in patient ER11 where TS (b) performed segmentations similar to delineator one (a). The DSCs and HD95 values for the small bowel, colon and bowel are plotted in Figure 4.9, illustrating high DSCs (> 0.80) for all three structures. The high DSCs in combination with the low HD95 values for the small bowel and bowel imply that TS managed to identify these structures with adequate performance in this slice. The high HD95 value for the colon, close to 40 mm, indicates that the algorithm did not perform equally precise segmentations for this structure in the current slice. Qualitative analysis of Figure 4.8 shows that TS failed to recognise the patient's ascending colon, which appears in the upper left of the CBCTs on the patient's right side.

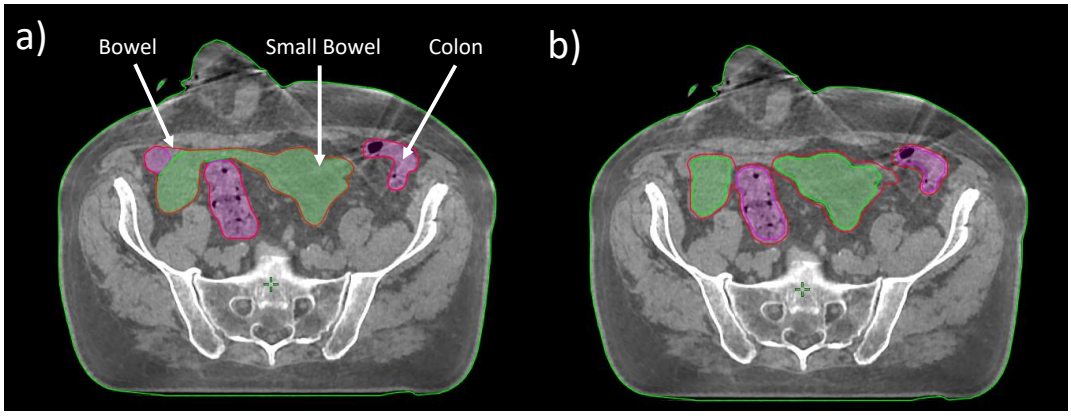

Figure 4.8: Illustration of a post-CBCT slice in fraction one of patient ER11 with the small bowel (green), colon (magenta) and bowel (red) segmentations by a) delineator one and b) TotalSegmentator, where TotalSegmentator performed satisfactory delineations.

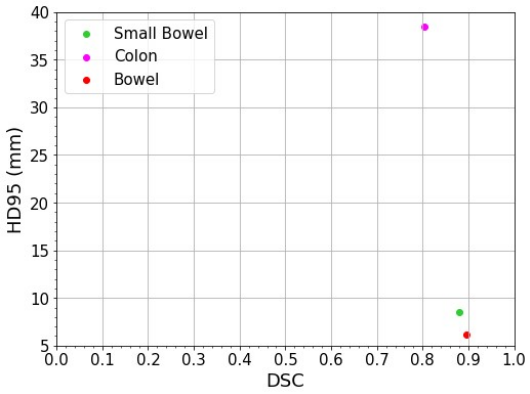

Figure 4.9: Scatter plot of 95th Hausdorff Distances (HD95s) against Dice Similarity Coefficients (DSCs) between delineator one's and TotalSegmentator's segmentations of the small bowel (green), colon (magenta) and bowel (red) in a post-CBCT slice in fraction one of patient ER11 where TotalSegmentator performed satisfactory segmentations.

A post-CBCT slice of patient ER1 where TS performed unsatisfactory segmentations is illustrated in Figure 4.10 b). For reference, delineator one's segmented bladder, small bowel, colon and bowel are shown in Figure 4.10 a). The DSCs and HD95 values comparing these segmentations are plotted in Figure 4.11. TS did not identify any small bowel correctly, if the segmentations by delineator one is considered as the benchmark. The DSC equalling zero for the small bowel is consistent with the observation of zero overlap between the segmented small bowels, and the HD95 value of approximately 32 mm suggests that the segmentations do not align. In the upper right of the CBCT in Figure 4.10 b), i.e. on the patient's left, there are air bubbles indicating bowel which TS ignored. Delineator one included this structure as colon, and hence also in the bowel segmentation, resulting in the mediocre DSCs for the colon and bowel. The colon resulted in a slightly higher HD95 value than the bowel because TS did not identify the patient's ascending colon, and incorrectly included some ascending colon as small

bowel. This incorrect definition has less impact on the bowel segmentation, where it is irrelevant whether the segmentations were initially defined as small bowel or colon. Therefore, the bowel achieves better DSCs and HD95 values than the small bowel and the colon, but still not satisfactory results. Furthermore, TS did not identify the bladder, and mistakenly included some small bowel within the bladder. Segmenting the bladder is generally straightforward, even in CBCTs. Figure 4.11 does not include geometrical comparison values for the bladder, since TS did not segment any bladder in the studied slice.

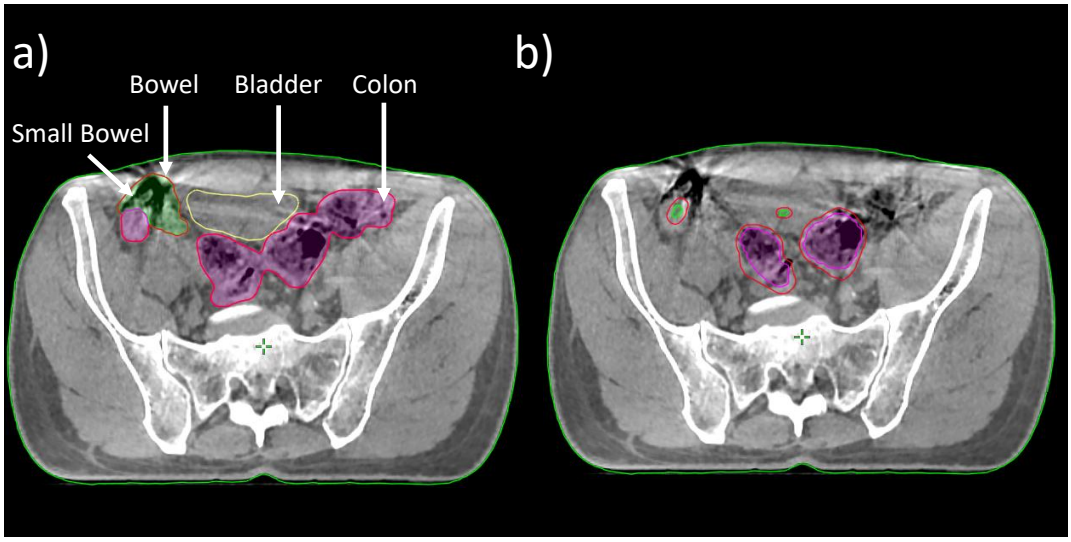

Figure 4.10: Illustration of one post-CBCT slice in patient ER1 with the bladder (yellow), small bowel (green), colon (magenta) and bowel (red) segmented by a) delineator one and b) TotalSegmentator, where TotalSegmentator had poor segmentation performance.

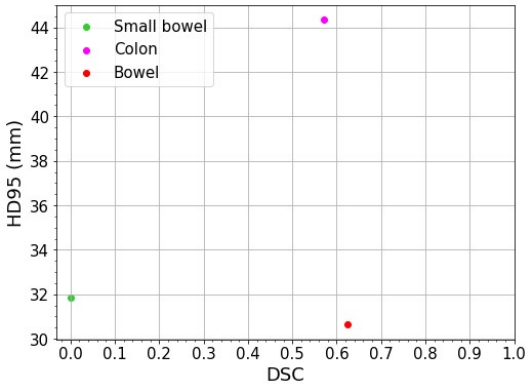

Figure 4.11: Scatter plot of 95th Hausdorff Distances (HD95s) against Dice Similarity Coefficients (DSCs) between delineator one's and TotalSegmentator's segmentations of the small bowel, colon and bowel in one post-CBCT slice of patient ER1, where TotalSegmentator had poor segmentation performance.

4.1.3 Dosimetric evaluation of AI segmentations

The dosimetric impacts of inaccurate segmentations were investigated because geometric evaluations do not reflect any clinical impact of the segmentations. Geometric inaccurate segmentations may give adequate measures of the dose distributions within the structures of interest. The Wilcoxon signed rank test showed that bowel V5 and V7.5 were statistically significantly greater for delineator one's segmentations than TS's segmentations ($p < 0.01$) when investigating segmentations in the pre-CBCTs of fx1 for all patients. Quantitative analyses of bowel V20 and V22.5 revealed that the volume of the bowels receiving at least 20.0 Gy and 22.5 Gy were slightly greater for TS's segmented bowels compared to the bowels segmented by delineator one, though with no statistically significant differences. DVHs based on the bowel volumes segmented by delineator one and TS are illustrated in Figure 4.12, showing the dose-volume-difference between the manually delineated and TS's segmented pre-treatment bowels of fx1. A positive volume difference implies that a greater volume of delineator one's than TS's segmented bowels received dose. For low doses, Figure 4.12 clearly illustrates that TS's segmentations in the majority of the patients underestimated the volume of irradiated bowels compared to delineator one's segmentations. For high doses, Figure 4.12 illustrates that the volume differences between the bowels segmented by delineator one and TS are small. The bowel V12.6, predictive of acute bowel toxicity, was not statistically significantly higher for segmentations by delineator one than TS (mean (range) 113.3 cc (30.8 cc-259.9 cc) and 109.0 cc (17.2 cc-246.4 cc), respectively).

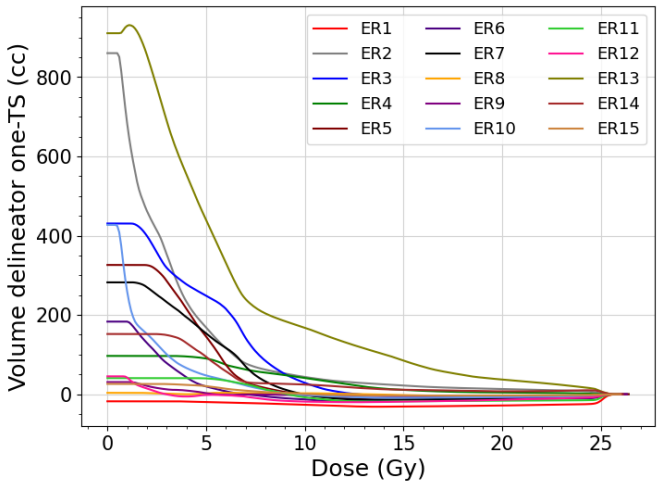

Figure 4.12: Dose-volume histogram for each patient representing the dose-volume-difference between delineator one's and TotalSegmentator's (AI) segmented pre-treatment bowels of fraction one. The y-axis represents the volume difference (cc) between the manually and automatically segmented bowels receiving a dose ≥ the dose (Gy) along the x-axis.

Patient ER13 stands out in Figure 4.12 as the patient with the greatest difference

between the volume of delineator one's and TS's bowels exposed to doses, for all dose levels. The CBCTs of ER13 were highly influenced by artifacts in all slices, which caused TS to exclude a great amount of the bowels in its segmentations, as illustrated in Figure 4.2. Hence, TS underestimated the dose exposed bowel volume. For patient ER1, the DVH illustrates that the irradiated volume of the bowel delineated by TS always were slightly greater than the volume segmented by delineator one, for all doses. In this patient, TS's bowel delineation always extended outside delineator one's bowel segmentation, affecting the estimated irradiated bowel volume for all dose levels, which can be caused by the post-processing.

4.2 Intra- and interfractional bowel motion

The Mann-Whitney U test showed that both the DSCs and HD95 values measuring intrafractional bowel motion (Figure 4.13 and 4.14, respectively) were statistically significantly higher and lower, respectively, than the DSCs and HD95 values measuring interfractional bowel motion (Figure 4.15 and 4.16, respectively) ($p < 0.01$). Hence, the interfractional bowel motion between fx1 and fx5 was significantly greater than the intrafractional bowel motion within both fx1 and fx5, as expected.

The box plots in Figure 4.13 and 4.14 show DSCs and HD95 values, respectively, measuring intrafractional bowel motion in each individual patient, comparing bowel loops in pre- and post-CBCTs of fx1 and fx5 separately. The measured intrafractional bowel motion varies between patients and between fractions for individual patients. For all patients, the median (Q1:Q3) DSC and HD95 value were similar for fx1 and fx5. In fx1, the results were 0.83 (0.76:0.88) and 12.59 mm (8.87:18.91 mm) respectively, comparable to the results in fx5 of 0.84 (0.76:0.89) and 13.00 mm (9.52:18.51 mm) respectively. Within some individual patients, the IQRs for DSCs and HD95 values in fx1 and fx5 are quite different in ranges. Therefore, the Mann-Whitney U test was used to investigate the difference in intrafractional bowel motion between fx1 and fx5, resulting in no statistically significant disparity. Observations revealed that the lowest DSCs and highest HD95 values causing the outliers in the intrafractional bowel motion plots (Figure 4.13 and 4.14) normally resulted from the structures in the most caudal slices. An intrafractional variation in the bowel's caudal extent can, for instance, be due to intrafractional changes in bladder filling. In the box plots, a larger box means a greater IQR and hence a greater spread of the data. Measuring intrafractional bowel motion, this means that there is a difference in the amplitude of the location shift of the bowel among the various CBCT slices in a single scan. In some slices, the pre- and post-treatment bowels were observed to be located close to each other, while in other

slices the differences between their positions were greater. A patient with a low median DSC and a high median HD95 value, combined or not combined with a wide IQR, has a bowel that moves measurably intrafractionally.

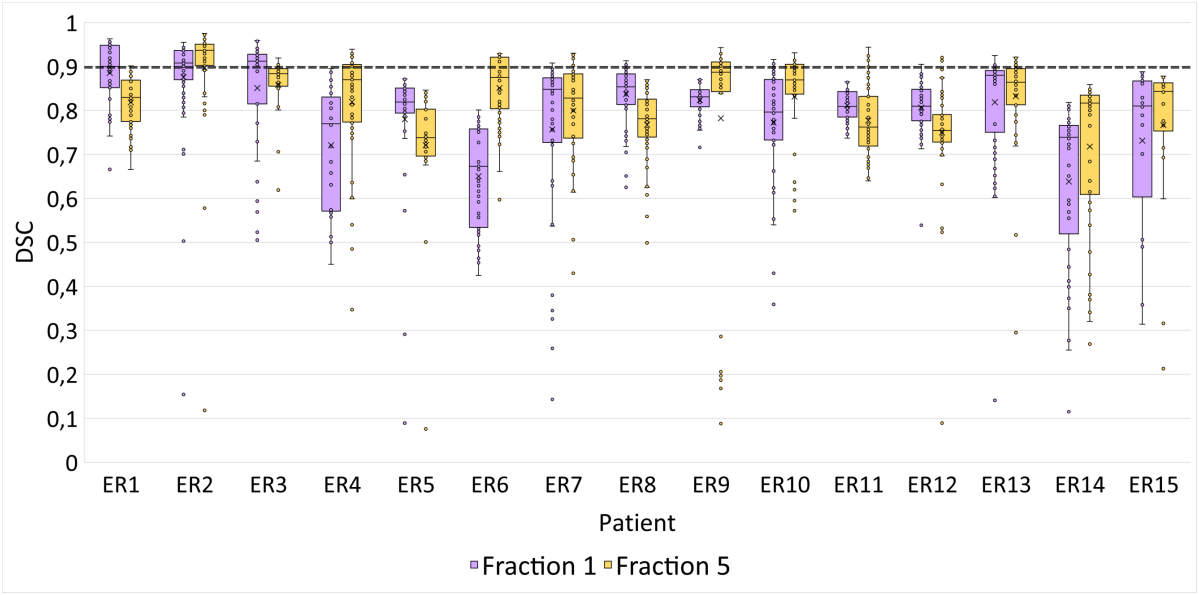

Figure 4.13: Intrafractional bowel motion in fraction one (purple) and fraction five (yellow). Box plots of Dice Similarity Coefficients (DSCs) for individual patients calculated between the segmented bowel loops in pre- and post-CBCTs. Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Whiskers are ranging from the minimum to the maximum value that is not an outlier. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3. The dashed line is the median DSC (0.90) for all patients between the bowel segmentations by delineator one and two, indicating segmentation uncertainties.

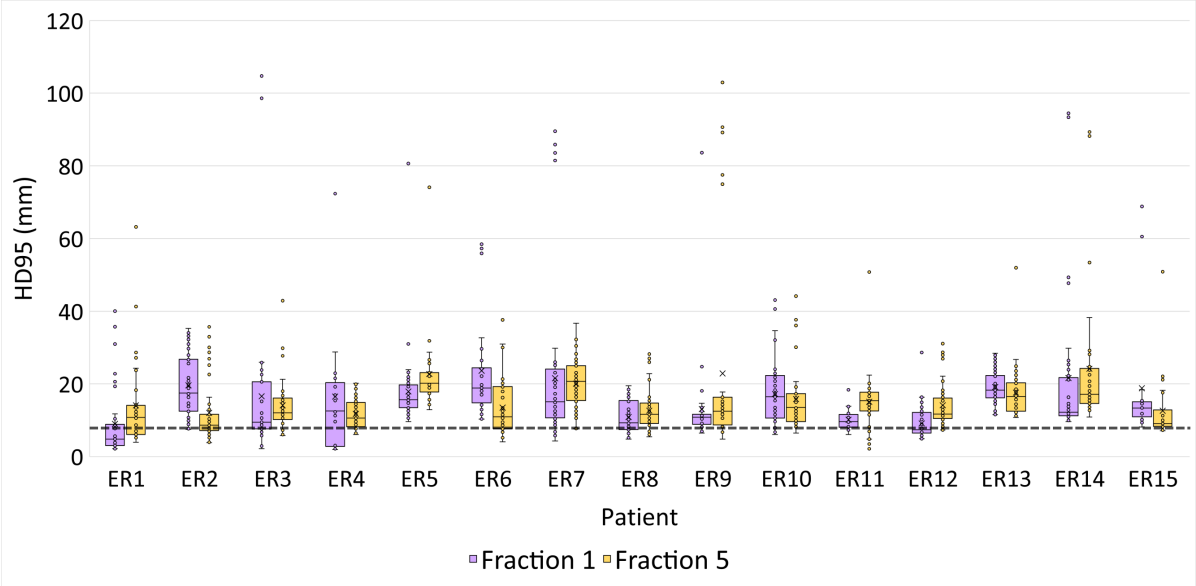

Figure 4.14: Intrafractional bowel motion in fraction one (purple) and fraction five (yellow). Box plots of 95th percent Hausdorff Distances (HD95s) for individual patients calculated between the segmented bowel loops in pre- and post-CBCTs. Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Whiskers are ranging from the minimum to the maximum value that is not an outlier. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3. The dashed line is approximately the median HD95 value (7.69 mm) for all patients between the bowel segmentations by delineator one and two, indicating segmentation uncertainties.

DSCs and HD95 values measuring interfractional bowel motion between the bowel loops in the pre-CBCTs of fx1 and fx5 are illustrated in Figure 4.15 and 4.16, respectively. The figures illustrate a great inter-patient variation in interfractional bowel motion. The box plots for some individual patients have a wide IQR, meaning that the location shift of the bowel between fx1 and fx5 varied among the CBCT slices. For all patients, the median (Q1:Q3) DSC and HD95 value measuring interfractional bowel motion were 0.72 (0.52:0.82) and 20.97 mm (15.39:31.79 mm), respectively.

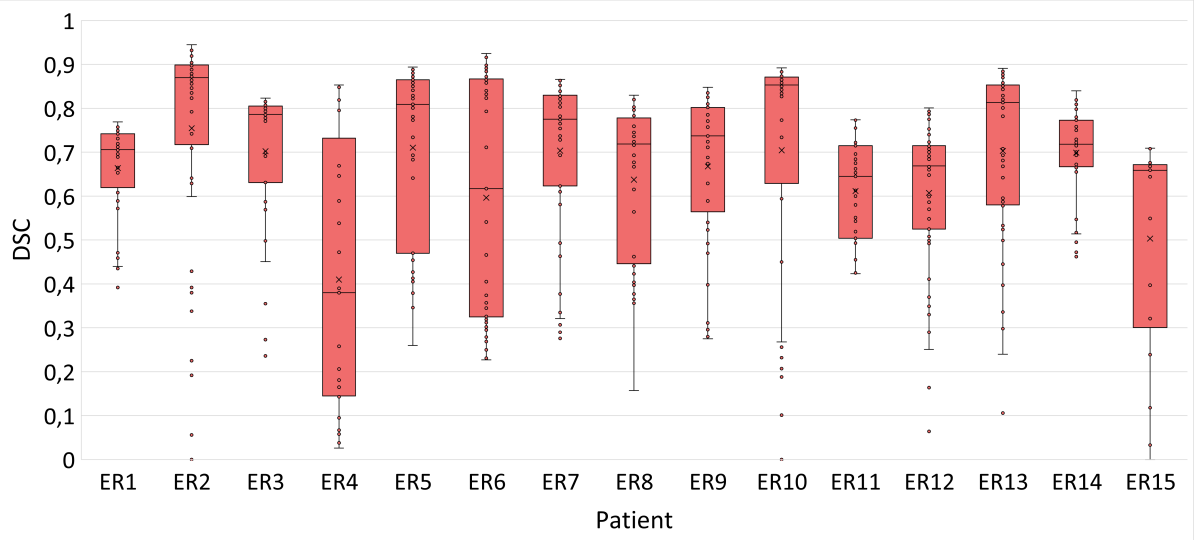

Figure 4.15: Interfractional bowel motion measured between fraction one and fraction five. Box plots of Dice Similarity Coefficients (DSCs) for individual patients calculated between the segmented bowel loops in pre-CBCTs of fraction one and fraction five. Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3.

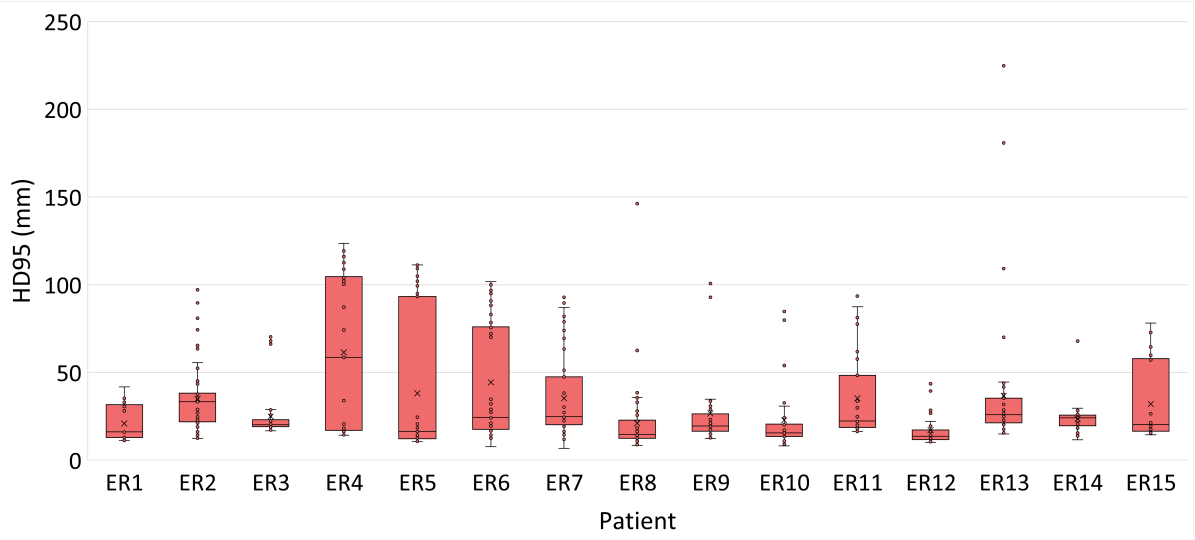

Figure 4.16: Interfractional bowel motion measured between fraction one and fraction five. Box plots of 95th percent Hausdorff Distances (HD95s) for individual patients calculated between the segmented bowel loops in pre-CBCTs of fraction one and fraction five. Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3.

Patient ER4 had the greatest interfractional bowel motion (median (Q1:Q3) DSC 0.38 (0.15:0.67), HD95 value 58.61 (17.12:104.53 mm)) among all the patients. The patient's bowel in fx5 had a much more caudally extent than in fx1 which, in addition to the bowel moving itself, is explained by the great difference in bladder volume,

illustrated in Figure 4.17. The bladder in fx5 was nearly empty while the bladder in fx1 was full (Table C6.2). This motivated us to look closer into the relation between bladder and bowel volume.

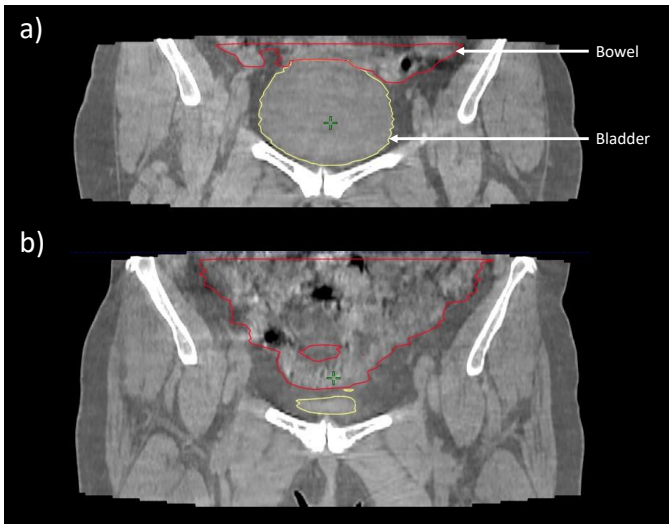

Figure 4.17: Pre-CBCT slices of patient ER4 with the bladder (yellow) and bowel (red) segmented in a) fraction one and b) fraction five, illustrating interfractional bowel motion and difference in bladder volume. In b), the bowel extends further inferior into the region previously occupied by the bladder in a).

4.3 Dosimetric evaluation

4.3.1 Impact of intrafractional bowel motion on bowel exposure

DVHs based on the pre- and post-treatment bowel loops suggested that the intrafractional bowel motion affected the dose distributions within the bowels for adaptive treatments in both fx1 and fx5, especially for low dose levels. The paired Wilcoxon signed rank test and Levene's test showed no statistically significant differences for the dose-volume parameters representing low doses (V5 and V7.5) and the parameters representing high doses (V20 and V22.5) between fx1 and fx5.

Figure 4.18 illustrates the DVHs for adaptive treatments for the dose-volume-differences between the manually delineated pre- and post-treatment bowels of fx1 and fx5 due to intrafractional bowel motion. Each line represents the volume difference for one patient between the irradiated volume of the pre- and post-treatment bowel for all dose levels. A positive volume difference means that a greater volume of the pre- than the post-treatment bowel received dose. As Figure 4.18 illustrates, for the majority of the patients, a greater volume of the pre- than the post-treatment bowels were

exposed to doses for almost all dose levels in both fractions. This result suggests that, in both fx1 and fx5, the intrafractional bowel motion moved parts of the bowels away from the pelvic regions exposed to radiation. The greatest volume differences are found between volumes exposed to low doses, approximately below 10-15 Gy. The bowels exposed to low doses are located in the cranial pelvic regions, where the segmentation uncertainties were highest due to artifacts. Inaccurate defined bowel volumes will influence the dose-volume-differences. The volume differences tend to decrease with increasing dose, i.e. for bowels located closer to the target and where the segmentation performances often were more accurate due to more visible bowel loops. Moreover, the volume differences appear to be more frequently negative for higher doses compared to lower doses.

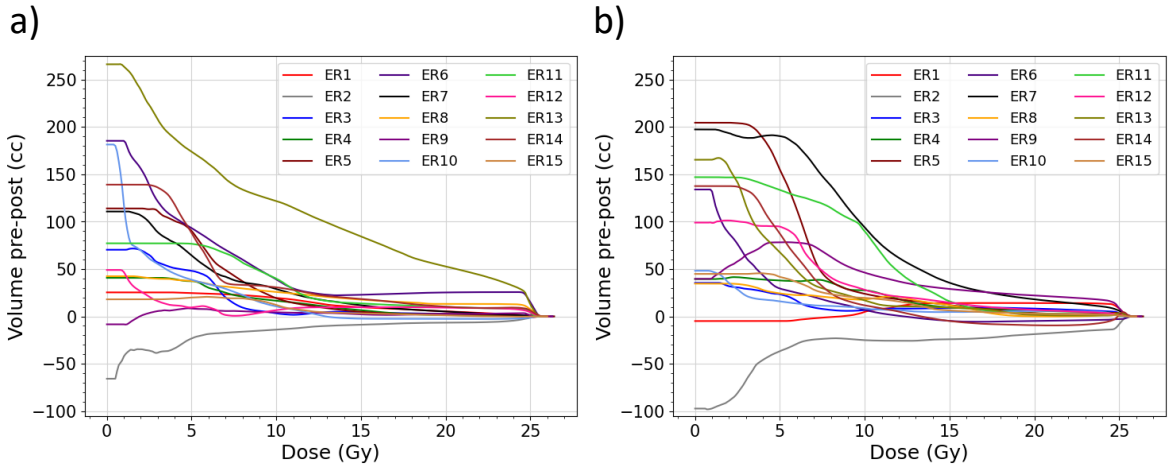

Figure 4.18: Dose-volume histograms (DVHs) for adaptive treatments for the dose-volume-differences between the manually delineated pre- and post-treatment bowels of fraction one (a) and fraction five (b) due to intrafractional bowel motion, for each individual patient. The y-axis represents the volume difference (cc) between the pre- and post-treatment bowels receiving a dose (Gy) $\geq$ the dose along the x-axis.

Due to uncertainties regarding the normality for the volume parameters V7.5, V20 and V22.5 (Figure A6.2), the paired Wilcoxon signed rank test and Levene's test were used to compare the distribution and variation between the dose-volume-difference parameters from fx1 and fx5. These tests were performed even though the individual value plot in Figure A6.3 suggested a relation between the parameters from fx1 and fx5 in some patients. For instance ER2 had the lowest dose-volume-difference parameters among all patients for all studied dose levels in both fractions, with consistently larger irradiated volumes in fx1 than in fx5. However, for most patients, it does not seem to be a relation between the individual patient's parameter in fx1 and fx5. Nor does the variation between the patients follow a clear pattern, and the variations between fx1 and fx5 seem to be approximately equal. For this reason, and since the paired Wilcoxon

signed rank test and Levene's test showed no statistically significant differences for any of the dose-volume-difference parameters between fx1 and fx5, the volume differences between the manually segmented pre- and post-treatment bowel loops for the DVH parameters V5, V7.5, V20 and V22.5 from fx1 and fx5 were gathered and plotted together in Figure 4.19. In the figure, low doses are represented by 5.0 Gy and 7.5 Gy (Figure 4.19 a) while high doses are represented by 20.0 Gy and 22.5 Gy (Figure 4.19 b). The box plots in these figures illustrate that the volume differences tend to be greater for lower doses, and the most negative differences occur for the higher dose levels, which agrees with the DVHs in Figure 4.18. The median values, represented by the middle line in each box, indicate a diversity among the volume differences for the chosen dose levels. At 5.0 Gy, 7.5 Gy, 20.0 Gy and 22.5 Gy, the volume differences in fx1 and fx5 combined have a median value of 40.16 cc, 27.78 cc, 4.30 cc and 2.81 cc, respectively. The greatest spread in the volume differences is observed for the volumes receiving $\geq$ 5 Gy, consistent with the DVHs where the spread of the dose-volume-difference histograms for different patients are most dispersed for lower doses and accumulates approximately around zero for higher doses.

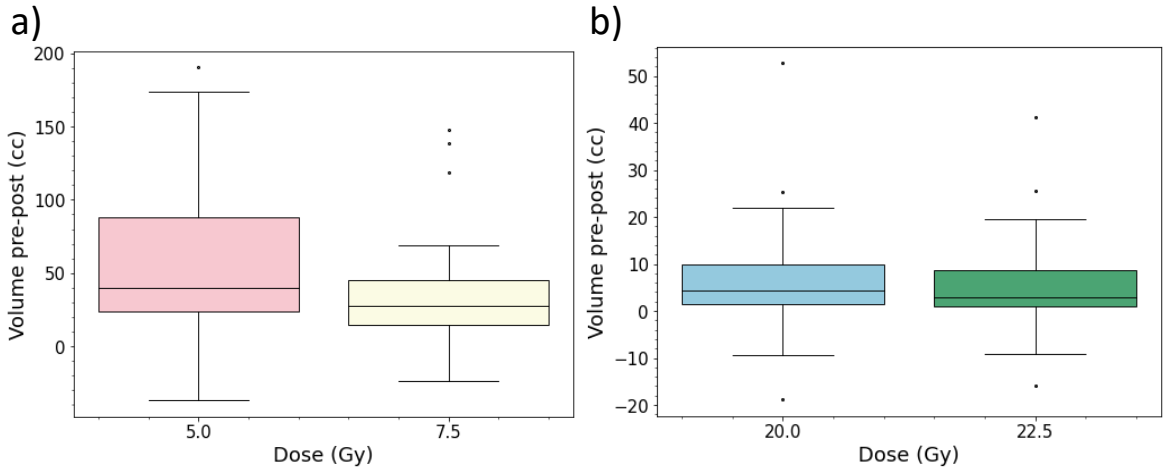

Figure 4.19: Box plots illustrating the volume differences (cc) between the manually segmented pre- and post-treatment bowels for the dose-volume histogram (DVH) parameters of the volumes receiving at least 5.0 Gy (V5), 7.5 Gy (V7.5), 20.0 Gy (V20) and 22.5 Gy (V22.5) for all patients in fraction one and fraction five. a) Volume differences due to intrafractional changes in bowel volumes exposed to low doses (V5 and V7.5). b) Volume differences due to intrafractional changes in bowel volumes exposed to high doses (V20 and V22.5). Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median and whiskers extending outside the box. Outliers are illustrated as the data points outside the whiskers, greater than 1.5 times the interquartile range (IQR) which is the distance between Q1 and Q3.

The intrafractional bowel motion in patient ER2 resulted in the worst dosimetric outcome in both fx1 and fx5 among all patients, illustrated in the DVH's in Figure 4.18. ER2's DVHs have the most negative volume differences for all doses, indicating that

the bowels moved into pelvic regions exposed to doses during both fx1 and fx5. The patient's bowel contained a great amount of air both during fx1 and fx5, as illustrated in Figure 4.20 a) and b) for fx5. Comparing the pre- and post-CBCTs illustrated in Figure 4.20 a) and b), respectively, shows that the air contained in the pre-treatment bowel moved in the caudally direction during RT delivery, resulting in the post-treatment bowel being located closer to the RT target, exposed to higher radiation doses. Patient ER13's DVH in fx1, illustrated in Figure 4.18 a), indicates that ER13 had the most dosimetric benefit caused by the intrafractional bowel motion. A pre- and post-CBCT slice of ER13 with bowels of high air contents are illustrated in Figure 4.20 c) and d), respectively. Qualitative analysis of the bowels superior to the bladders implies that the intrafractional bowel motion and bladder filling caused the bowel to move in the cranial direction during RT delivery, positioning the post-treatment bowel further away from the RT target than the pre-treatment bowel. However, both ER2 and ER13 were among the patients with the greatest amount of artifacts in their CBCTs (Figure 4.2 for ER13), resulting in uncertain manual bowel segmentations which may have influenced the volume differences, especially for the cranial bowels exposed to low doses.

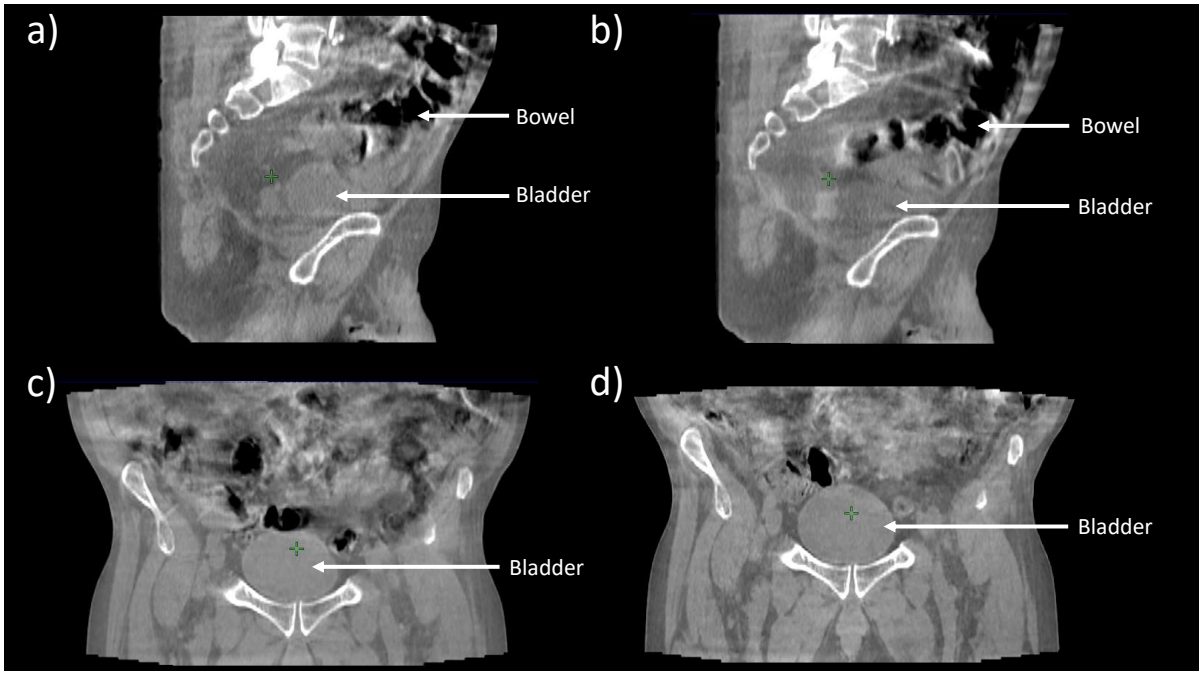

Figure 4.20: CBCTs explaining the outliers (patient ER2 and ER13) in the dose-volume histograms in Figure 4.18. a) Pre- and b) post-CBCT slice of patient ER2 in fraction five illustrating an intrafractional displacement of the bowel due to displacement of the air bubbles. c) Pre- and d) post-CBCT slice of patient ER13 in fraction one illustrating intrafractional displacement of the bowel superior to the bladder in the cranial direction due to bowel motion and intrafractional bladder filling.

In fx1, patient ER10's DVH illustrates a rapid decrease for increasing doses below approximately 3.0 Gy. The pre-CBCTs were more degraded by artifacts than the post-CBCTs, especially in the cranial slices where the bowels receiving the lowest doses are

located. Therefore, the pre-treatment bowel was segmented resembling a bowel bag structure, while the post-treatment bowel was defined as individual loops, as illustrated in Figure 4.21. Since the bowel bag contained a greater total segmented volume than the bowel loop segmentation, this gave the false impression that the volume of the irradiated pre-treatment bowel was greater than the volume of the irradiated post-treatment bowel. In the more caudal slices containing the bowels exposed to higher doses, less artifacts were present in the CBCTs, and the individual bowel loops could be segmented in both the pre- and post-CBCTs. Hence, the rapid decrease in the volume differences the more caudal the bowels were located, closer to the RT target.

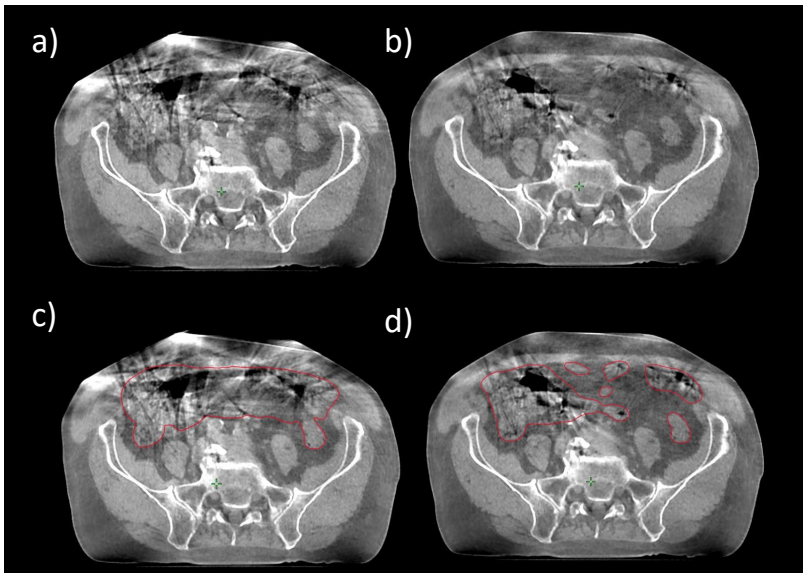

Figure 4.21: CBCT slices of patient ER10 in fraction one. a) Pre-CBCT slice with poor image quality due to artifacts, cranially located within the pelvis. b) Post-CBCT of the corresponding slice as illustrated in a), less occupied by artifacts. c) The bowel illustrated in a) manually segmented similar to a bowel bag structure. d) The bowel loops illustrated in b) manually segmented.

Patient ER6 is a typical patient regarding intra- and interfractional bowel motion (figs. 4.13 to 4.16). However, observations revealed that great parts of its colon loop was located close to the RT target in all fractions, as illustrated in Figure 4.24 for fx1 and fx5, exposed to high doses in both fractions, shown in the DVHs in Figure 4.22. Thus, the patient is an example of a "worst case" patient regarding bowel exposure. The patient had greater intrafractional motion in fx1 than fx5, with a median (Q1:Q3) DSC and HD95 value of 0.67 (0.55:0.76) and 18.89 mm (15.91:24.28 mm) respectively in fx1, and 0.88 (0.81:0.92) and 10.93 mm (7.75:18.98 mm) respectively in fx5. For interfractional bowel motion between fx1 and fx5, the DSC was 0.62 (0.33:0.87) and the HD95 value was 24.34 mm (17.69:75.80 mm). Qualitative analyses of the pre- and post-CBCTs of fx1 illustrated an intrafractional difference in the bowel's position superior to the bladder due to intrafractional bladder filling, illustrated in Figure 4.23

a) and b). Additionally, the air observed in the most caudal pre-treatment bowel loops moved in the cranial direction during treatment. In fx5, an equally large intrafractional displacement of air as in fx1 was not observed. A great interfractional displacement of the bowel loops positioned superior to the bladder was observed due to a significantly larger bladder volume in fx1 compared to fx5 (Table C6.2), illustrated in Figure 4.23 c) and d). Therefore, the intra- and interfractional bowel motion mostly displaced the bowels located cranially relative to the RT target. Qualitative analyses of the CBCTs were important to detect the fixed colon loop that did not move out of high dose regions during. In Figure 4.24 a) and c), a segment of the post-treatment bowel (colon) loop in fx1 and fx5, respectively, are illustrated together with the CTVs. The dose distributions in Figure 4.24 b) and d) show that parts of the bowel loop in both fractions receives high doses.

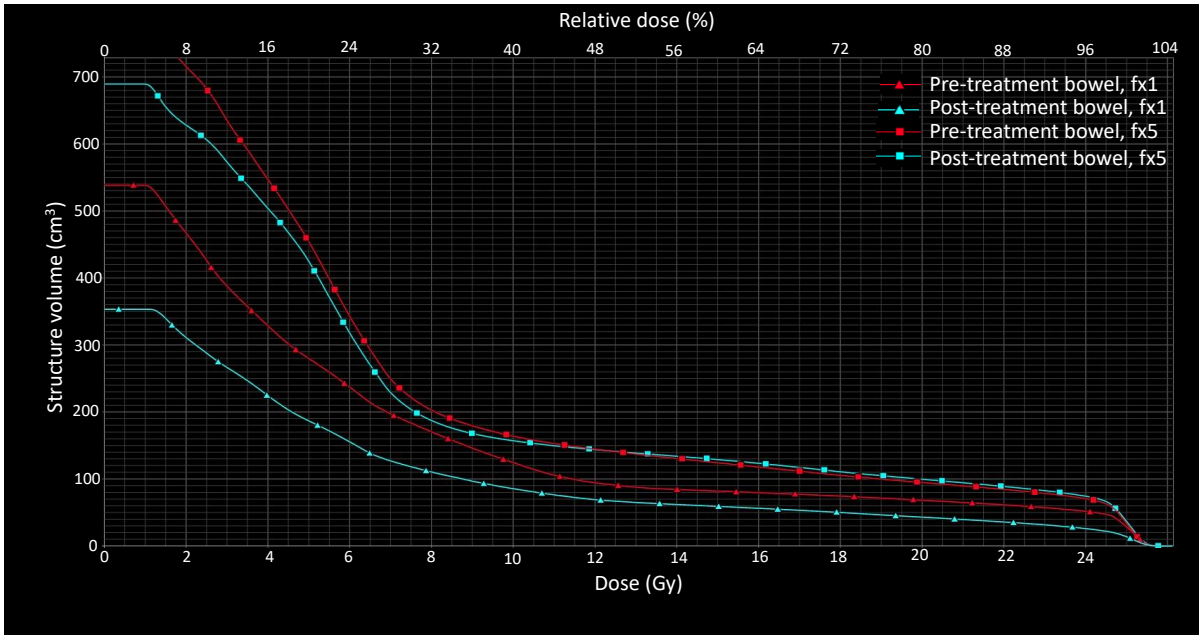

Figure 4.22: Dose-volume histograms (DVHs) for patient ER6's manually segmented pre- and post-treatment bowel loops in fraction one (fx1) and fraction five (fx5).

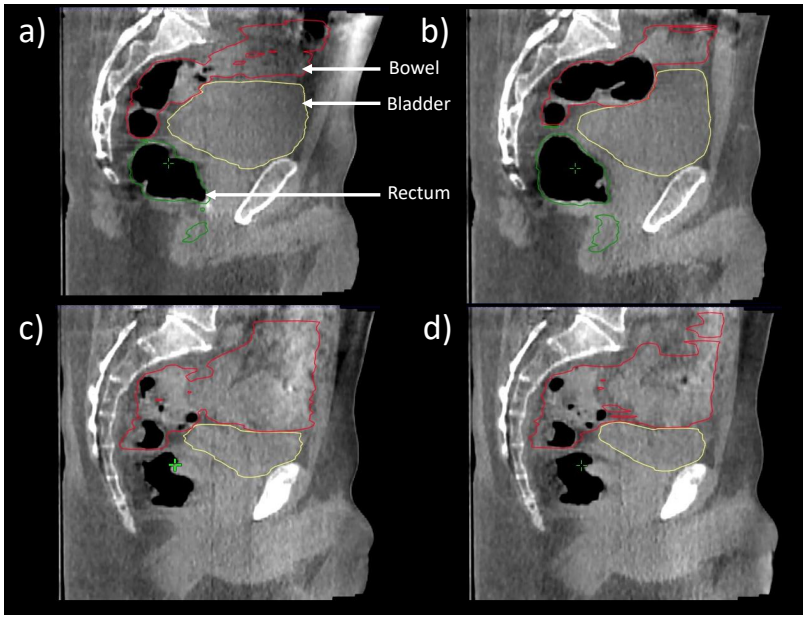

Figure 4.23: Sagittal CBCT slices of patient ER6 illustrating intra- and interfractional bowel motion and changes in the bladder volume. a) Pre-CBCT of fraction one (fx1), b) post-CBCT of fx1, c) pre-CBCT of fraction five (fx5) and d) post-CBCT of fx5.

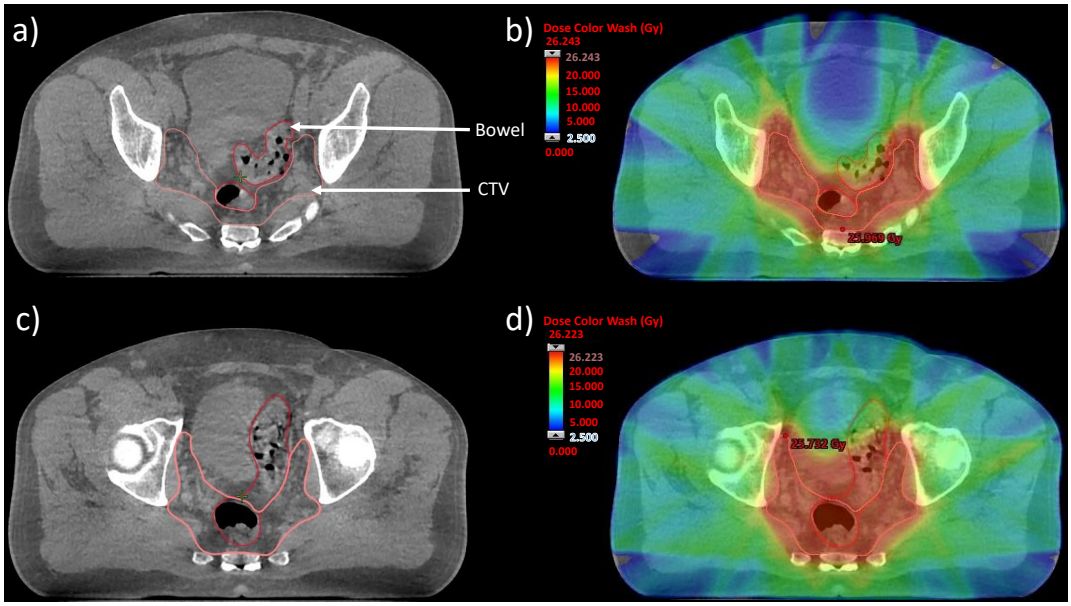

Figure 4.24: Post-CBCT slices of patient ER6 illustrating a fixed bowel loop located close to the CTV in a) fraction one (fx1) and b) fraction five (fx5). The simulated dose distributions for the post-CBCTs in fx1 and fx5 are illustrated in b) and d), respectively, with a color wash where red regions are exposed to high doses and blue regions are exposed to low doses according to the color bars.

ART may be appropriate when a fixed bowel loop is permanently located close to the target, always exposed to high doses, as in the example of ER6. Also, patients ER1, ER5, ER8, ER12, ER13 and ER14 had a fixed bowel loop permanently located close to the RT target in all fractions, in addition to the bowel that has to be close to the target due to the continuous transition between the rectum and colon. In ART, the treatment

can adapt to the bowel loop that is likely to retain in the same position during the RT delivery fraction to minimize the dose delivered to the loop.

4.3.2 Impact of bladder filling on bowel exposure

Due to observations made when investigating the intra- and interfractional bowel motions and the dosimetric impact of bowel motion, we were motivated to examine the relations between bladder volume, bowel displacement and bowel exposure. For the pre-treatment bladders and bowels in both fx1 and fx5, Figure 4.25 a) illustrates a slight relation between bladder volumes and bowel volumes exposed to low doses. Patients with full bladders (volume $\geq$ 300 cc) seem to have less bowel V10s compared to patients with less bladder volumes (volume $<$ 300 cc). The t-test did not indicate statistically significant differences for bowel volumes exposed to at least 10.0 Gy and 20.0 Gy between patients with full pre-treatment bladders and patients with smaller bladder volumes both in fx1 and fx5. A general relation between bladder volumes and irradiated bowel volumes in Figure 4.25 is difficult to detect for both dose levels, and the strengths of the hypothesis tests are weakened by the small number of patients in our study, which was insufficient to reveal trends in our data. Four patients in fx1 and one patient in fx5 had pre-treatment bladder volumes $\geq$ 300 cc. Since only one patient (ER1) had a pre-treatment bladder volume $\geq$ 300 cc in fx5, nothing can be said about the distribution of the bowel V10s or V20s, and they could not be tested by the Shapiro-Wilk test or the F-test. The F-test between the pre-treatment bowel V10s in patients with full and less full bladders in fx1 indicated no significant difference between their variances. The same applied to the pre-treatment bowel V20s in fx1.

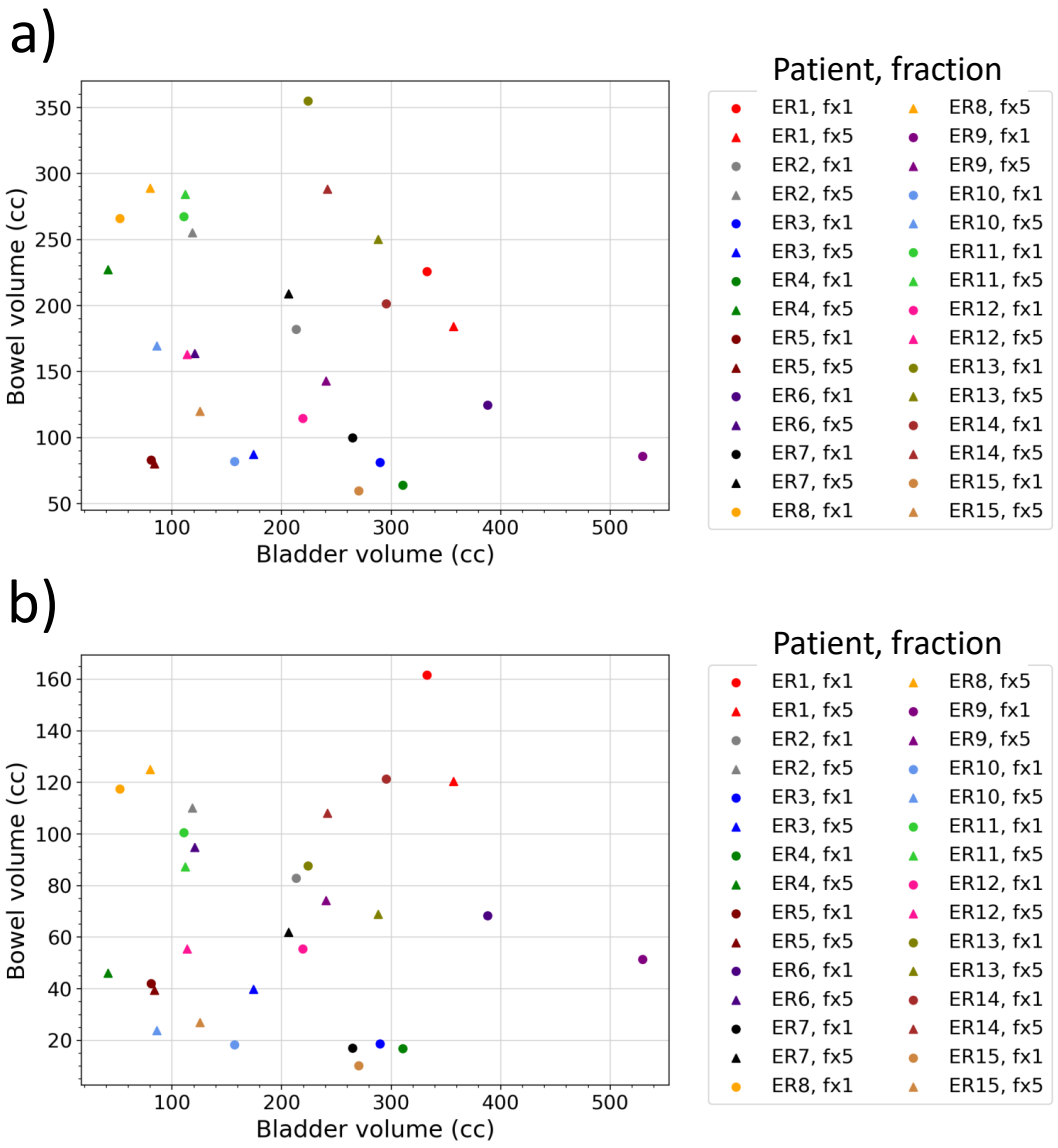

Figure 4.25: The pre-treatment bladder volumes (cc) plotted against the manually segmented pre-treatment bowel volumes (cc) exposed to least a) 10.0 Gy (V10) and b) 20.0 Gy (V20) in fraction one (fx1) and fraction five (fx5) for all patients.

Patient ER13 did not have the smallest pre-treatment bladder volume in fx1 among all the patients. Nevertheless, ER13 in fx1 had the greatest volume of the pre-treatment bowels receiving at least 10.0 Gy (Figure 4.25), consistent with the patient's DVH in fx1 (Figure 4.18 a), which may be caused by segmentation uncertainties due to artifacts, as explained in section 4.3.1.

Table C6.2 gives an overview of the pre- and post-treatment bladder volumes for each patient in fx1 and fx5. The relative intrafractional changes in bladder volumes are given, where a positive relative change means that the bladder volume has increased during the fraction relative to the bladder volume before the RT delivery, and vice versa for a negative change. Regarding both pre- and post-treatment bladders in all patients,

there was a large difference between the maximum and minimum bladder volume in both fx1 and fx5. The mean (range) absolute intrafractional difference in bladder volumes for all patients were similar in fx1 and fx5, measuring −10.9 cc (−48.9-16.1 cc) and −6.4 cc (−53.9-21.5 cc), respectively. The absolute interfractional difference between the pre-CBCTs in fx1 and fx5 were also similar to the interfractional difference between the post-CBCTs, measuring 89.9 cc (−64.1-288.8 cc) and 94.4 cc (−36.4-314.0 cc), respectively. A negative intrafractional difference means that the volume of the post-treatment bladder was greater than the pre-treatment bladder, and a negative interfractional difference means that the pre- or post-treatment bladder in fx5 had greater volume than the pre- or post-treatment bladder in fx1. Small intra- or interfractional bladder volume differences in individual patients may be due to segmentation uncertainties. In fx1, the patients ER5 and ER14 had the greatest relative intrafractional increases in bladder volumes among all patients (17%). ER5 also had the greatest relative bladder volume increase in fx5 (31%). Pre- and post-CBCTs of ER14 in fx1 are illustrated in Figure 4.26. The pre-treatment bladder in Figure 4.26 a) and 4.26 c) has a reduced volume and less cranial extent than the post-treatment bladder in Figure 4.26 b) and 4.26 d). The intrafractional increase in the bladder volume resulted in the bowel being lifted in the cranial direction, further away from the rectum, as illustrated in the CBCTs.

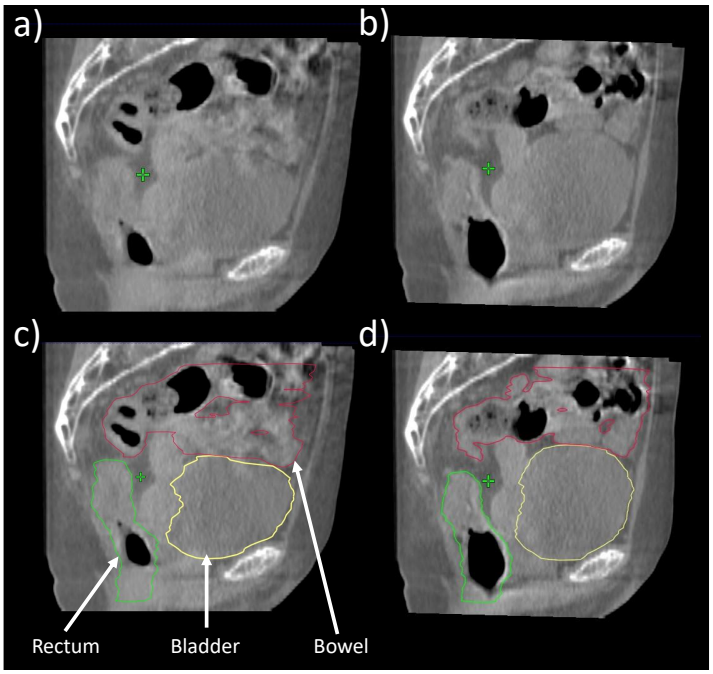

Figure 4.26: A sagittal a) pre-CBCT and b) post-CBCT slice of patient ER14 in fraction one illustrating an intrafractional increase in bladder volume. c) The bladder (yellow), bowel (red) and rectum (green) illustrated in a) are manually segmented. d) The bladder, bowel and rectum illustrated in b) are manually segmented. The intrafractional increase in bladder volume results in an intrafractional displacement of the bowel in the cranial direction.

Coefficients for the linear regression analyses and the Pearson correlation coefficients with p-values for relations between the pre-treatment bladder and irradiated bowel volumes and between the intrafractional differences in bladder and bowel volumes for all patients in both fx1 and fx5 are given in Table 4.2. A weak negative, but statistically significant correlation was found between the intrafractional differences in bladder volumes and the intrafractional differences in bowel volumes exposed to at least 20.0 Gy, in both fractions. Hence, an intrafractional increase in the bladder volume, where the post-treatment bladder volume is greater than the pre-treatment bladder volume, implies an intrafractional decrease the bowel volume exposed to at least 20.0 Gy, where a greater volume of the pre-treatment bowel than the post-treatment bowel is exposed to dose.

Table 4.2: Coefficients for the linear regression analyses for the linear relationships between the pre-treatment bladder volumes and pre-treatment bowel volumes exposed to at least 10.0 Gy (V10) and 20.0 Gy (V20) in fraction one and five for all patients. The Pearson correlation coefficients (r) with p-values are given for the same analyses. Also, coefficients for the linear regressions and the Pearson correlation coefficients with p-values are given for the comparisons between the intrafractional changes in bladder volumes and the intrafractional differences in bowel volumes exposed to at least 10.0 Gy and 20.0 Gy in fraction one and five, for all patients.

Analyses of pre-treatment bladder volumes vs. pre-treatment irradiated bowel volumes				
	Intercept (β_0)	**Slope (β_1)**	**Correlation coefficient (r)**	**p-value**
Bowel V10	212.61	−0.19	−0.27	0.16
Bowel V20	66.94	0.01	0.02	0.91
Analyses of intrafractional changes in bladder volumes vs. intrafractional changes in bowel exposure				
	Intercept (β_0)	**Slope (β_1)**	**Correlation coefficient (r)**	**p-value**
Bowel V10	21.88	−0.33	−0.21	0.26
Bowel V20	4.31	−0.24	−0.38	0.04

The relations between an intrafractional difference in bladder volume and the volume difference between the pre- and post-treatment bowels exposed to at least 10.0 Gy and 20.0 Gy in both fx1 and fx5 are illustrated in Figure 4.27 together with the linear regression lines explaining the linear relationships. All patients experiencing an intrafractional increase in bladder volume had less volumes of their post-treatment bowels exposed to at least 10.0 Gy than their pre-treatment bowels in both fx1 ans fx5. The same applied to the volume differences between the pre- and post-treatment bowel V20s, except for one patient (ER14). In Figure 4.27, a positive change in bladder volume means that the bladder volume has decreased during the fraction, and opposite for a negative change. A positive bowel V10 or V20 difference means that the bowel exposure decreased during the fraction, and vice versa for a negative difference. Figure 4.27 illustrates a slight relation between an intrafractional increase in bladder volume and

decrease in bowel exposure, however this correlation was only statistically significant for bowel V20. In Figure 4.27 intrafractional decreases in bowel V10s and V20s were also observed in multiple patients with an intrafractional decrease in the bladder volume. A post-treatment bladder of reduced volume compared to the pre-treatment bladder volume is explained by segmentation uncertainties where the pre-treatment bladder is incorrectly segmented too large and/or the post-treatment bladder is segmented as a too small structure. Patients ER2, ER6 and ER8 had the most negative relative bladder volume differences (Table C6.2). These differences are caused by segmentation uncertainties due to bladders located close to bowels containing air, causing artifacts, or due to poor image quality. Patient ER2 had a measured intrafractional decrease in bladder volume in both fx1 and fx5, and greater volumes of its post-treatment bowel received at least 10.0 Gy and 20.0 Gy than its pre-treatment bowel. It is likely that some bowel has mistakenly been segmented as bladder in the post-CBCT, or vice versa for the pre-CBCT, causing the patient to be located in the lower right of the plots in Figure 4.27.

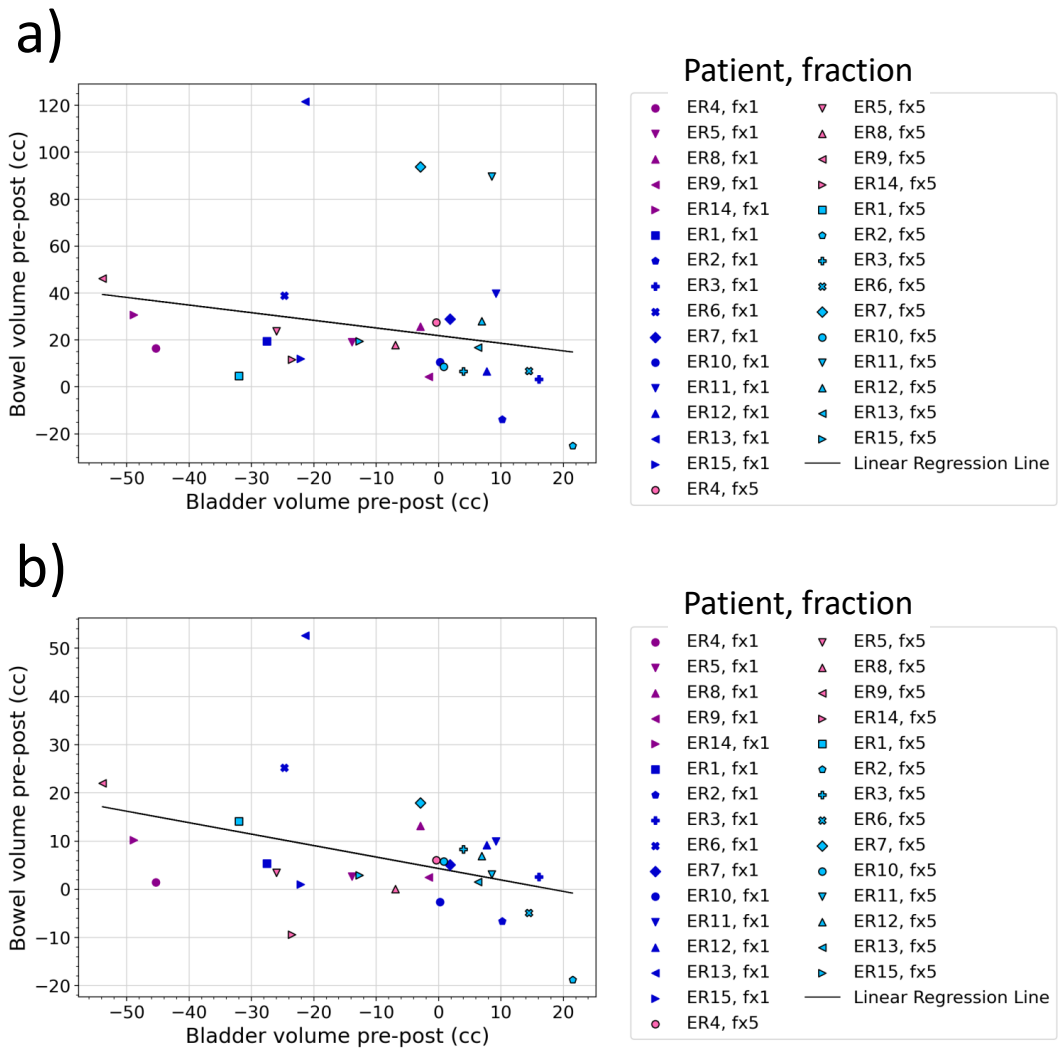

Figure 4.27: Absolute differences between the pre- and post-treatment bladder volumes plotted against the volume differences between the manually segmented pre- and post-treatment bowels receiving at least a) 10.0 Gy (V10) and b) 20.0 Gy (V20) in fraction one (fx1) and fraction five (fx5) for all patients. The linear regression lines representing the linear relationships between intrafractional bladder and bowel volume changes are also plotted. Male patients are plotted in dark blue for fx1 and light blue for fx5. Female patients are plotted in purple for fx1 and pink for fx5.

ER14's DSCs and HD95 values (Figure 4.13 and 4.14) indicate intrafractional bowel motion in both fx1 and fx5, explained by the bowel moving itself and changes in the bladder volume. The patient's median (Q1:Q3) DSC and HD95 value in fx5 were 0.82 (0.62:0.83) and 17.15 mm (14.81:24.11 mm) respectively. In both fx1 and fx5, ER14 experienced an intrafractional increase in bladder volume (Table C6.2). The patient's DVHs in Figure 4.18 illustrate that a greater volume of the pre-treatment bowels than the post-treatment bowels received low doses. As observed in the CBCTs in Figure 4.26, an intrafractional increase in the bladder volume forced the bowels to move in the cranial direction, away from the rectum. Therefore, it was expected that the

post-treatment bowels received less low doses than the pre-treatment bowels, which the DVHs support. In fx5, a greater volume of ER14's post-treatment bowel was exposed to doses above approximately 15.0 Gy than the pre-treatment bowel (DVH in Figure 4.18 b). A sagittal pre-CBCT slice of ER14 from fx5 is illustrated in Figure 4.28, where a small intrafractional increase in the bladder volume resulted in the post-treatment bowel to be shifted in the cranial direction compared to the pre-treatment bowel. This is evident in the sCT in Figure 4.28 c) where the delineated post-treatment bowel (blue) ends superior to the pre-treatment bowel (red). However, the bladder filling had low influence on the bowel volume exposed to high doses. Further, illustrated by image slices towards the patient's right side, the intrafractional displacement of the air within the bowel influenced the bowel to move into the high dose regions during RT delivery, regardless of the change in bladder volume.

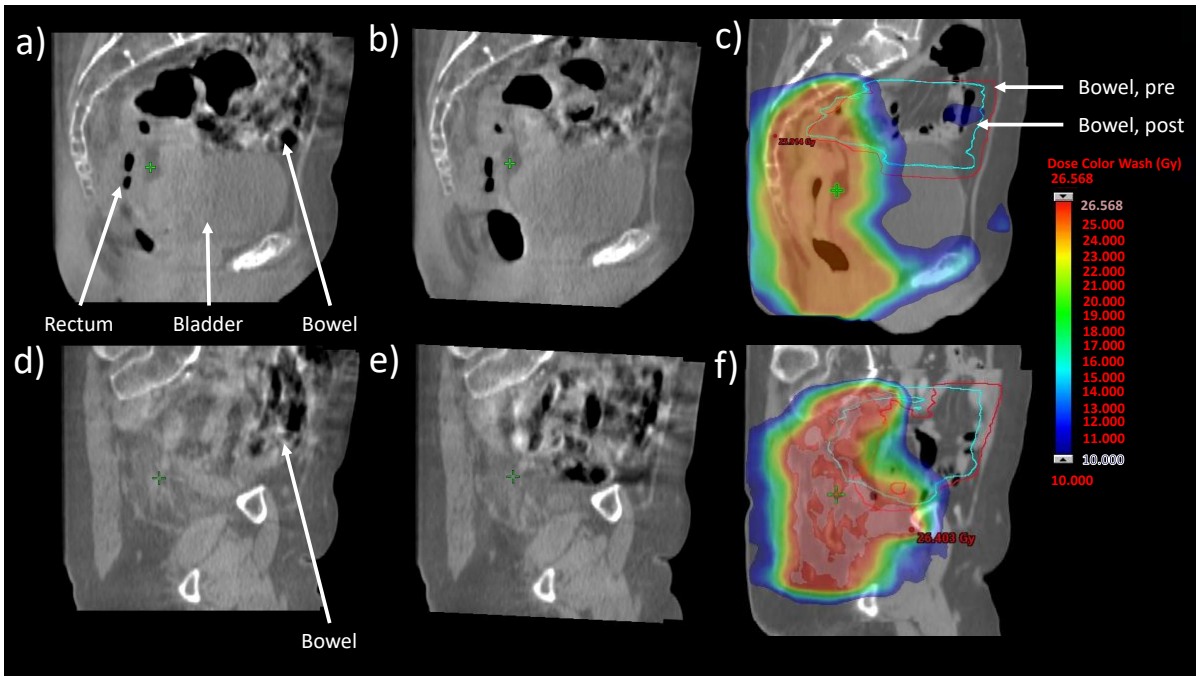

Figure 4.28: Pre- and post-CBCT slices in fraction five (fx5) and a synthetic CT (sCT) slice of patient ER14 illustrating differences in irradiated bowel volumes. a) and d) illustrate pre-CBCTs, b) and c) illustrate post-CBCTs. The slices in figure a), b) and c) correspond and are approximately the middle slices in the sagittal planes. The slices in figure d), e) and f) correspond and are located on the patient's right side. Figure c) and f) illustrate sCTs with simulated dose distributions, showing doses in the range from 10.000 Gy (dark blue) to 26.568 Gy (dark red). In Figure c), the manually delineated red and blue bowels are the pre-treatment bowel illustrated in a) and the post-treatment bowel illustrated in b) respectively. In Figure f), the manually delineated red and blue bowels are the pre-treatment bowel illustrated in d) and the post-treatment bowel illustrated in e) respectively.

4.4 Predicting grade ≥ 3 acute toxicity in bowel loops

Equation 4.1 gives the logistic regression curve relating small bowel V15 to predicted probability of grade ≥ 3 acute small bowel toxicity, shown by Banerjee et al. [92].

$$p(x) = \frac{e^{5.56+0.01x}}{1+e^{-5.56+0.01x}} \tag{4.1}$$

A dose of 12.6 Gy in 5 Gy fractions using SCRT is equivalent to a dose of 15 Gy in 1.8 Gy fractions using LCRT. For all patients, the median (Q1:Q3) predicted probability of grade ≥ 3 acute bowel toxicity for the V12.6 bowel loops were similar for the pre- and post-treatment bowels within the same fraction and between fx1 and fx5. The predicted probability was 0.01 (0.01:0.03) for the pre-treatment bowel loops in fx1, 0.01 (0.01:0.03) for the post-treatment bowel loops in fx1, 0.02 (0.01:0.04) for the pre-treatment bowel loops in fx5 and 0.01 (0.01:0.03) for the post-treatment bowel loops in fx5. The intrafractional differences in toxicity risk estimates were also similar in fx1 and fx5, resulting in 0.00 (0.00:0.01) for the intrafractional differences in toxicity risks in fx1 and 0.00 (0.00:0.01) in fx5. No great differences were observed for the interfractional differences in toxicity risks between fx1 and fx5 for the pre- and post-treatment bowels, which was -0.01 (-0.01:0.00) between the pre-treatment bowel loops in fx1 and fx5, and 0.00 (-0.01:0.00) between the post-treatment bowel loops in fx1 and fx5. A positive intrafractional difference in toxicity risk means that the predicted toxicity probability was greater for the irradiated pre-treatment bowel than the post-treatment bowel. A positive interfractional difference in toxicity risk means that the predicted toxicity probability was greater for the irradiated bowel in fx1 than the irradiated bowel in fx5. In total, 11 pre- and post-treatment bowels in fx1 and fx5 (five patients) had a predicted probability $\geq 5\%$ of grade ≥ 3 acute bowel toxicity, and no patients had predicted probability $\geq 10\%$. One patient (ER13) had an intra- and interfractional difference in toxicity risk $\geq 5\%$ (7% and 6% respectively).

The logistic regression curve and the predicted probabilities of grade ≥ 3 acute bowel toxicity for the pre- and post-treatment bowel loops receiving at least 12.6 Gy for each patient in fx1 and fx5 are illustrated in Figure 4.29. Figure 4.30 shows an enlarged version of the logistic regression curve and the predicted probabilities for the area where the predictions were located for a) patients ER1-7 and b) patients ER8-15. As the figures illustrate, five patients (ER1, ER2, ER8, ER13 and ER14) stand out with higher predicted toxicity risks than the other patients. These patients, except ER2, had fixed bowel loops permanently located close to the RT target in all fractions, as earlier explained. The pre- and post-treatment bowels of patient ER8 in fx1 and fx5 resulted in some of the greatest observed predicted toxicity risks, which all had predicted proba-

bilities $\geq 5\%$. Among all investigated bowels, the post-treatment bowel of patient ER8 in fx5 and the pre-treatment bowel of patient ER13 in fx1 had the greatest predicted toxicity risks (9%).

As earlier explained, patient ER6 is an example patient with a bowel loop receiving high doses in all fractions (Figure 4.24). ER6 did not result in high predicted probabilities of grade ≥ 3 acute bowel toxicity, explained by the patient's cranial bowels being shielded from irradiation due to a dorsal target volume with a cranial extension to the caudal S1 (Figure 2.17). Normally, the cranial extension of the target volume reaches the sacral promontory. Hence, the caudal bowel loops were exposed to high doses, however the volume of the irradiated cranial loops were reduced, resulting in lower toxicity estimates.

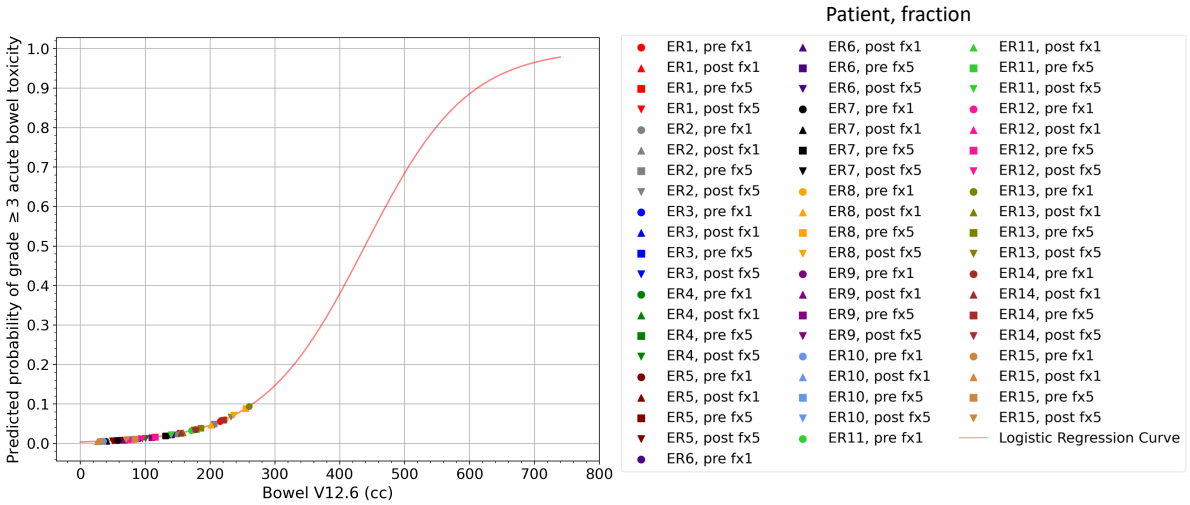

Figure 4.29: Logistic regression curve and predicted probabilities for grade ≥ 3 acute toxicities for bowel loops based on the bowel volume receiving at least 12.6 Gy (V12.6). Probabilities are predicted for pre- and post-treatment bowel loops in fraction one (fx1) and fraction five (fx5) for all 15 patients.

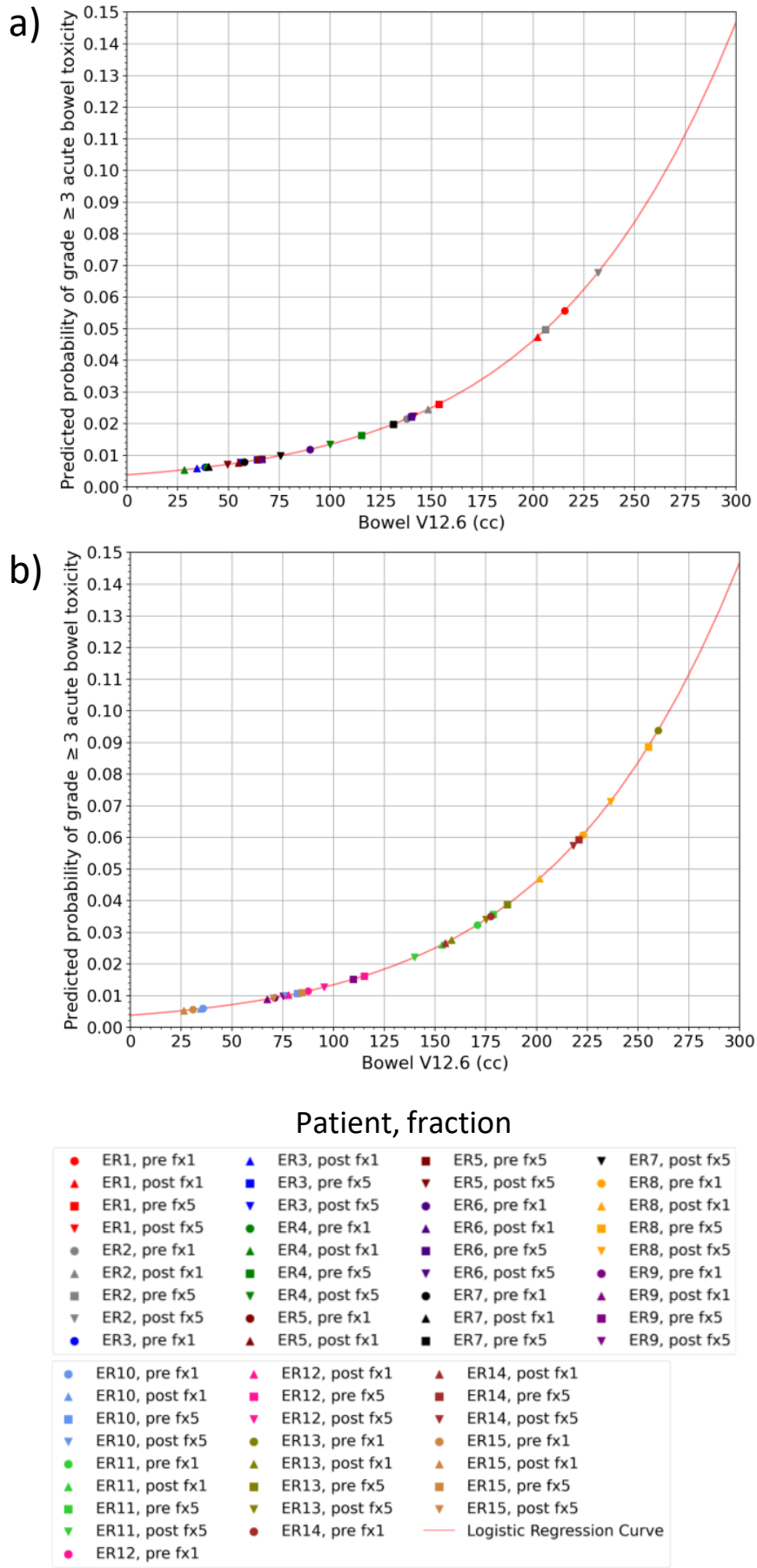

Figure 4.30: Logistic regression curve and predicted probabilities for grade ≥ 3 acute toxicities for bowel loops based on the bowel volume receiving at least 12.6 Gy (V12.6), illustrated for probabilities ≤ 0.15 and V12.6 ≤ 300 cc. Probabilities are predicted for pre- and post-treatment bowel loops in fraction one (fx1) and fraction five (fx5) for patients a) ER1-7 and b) ER8-15.

The probabilities of grade ≥ 3 acute bowel toxicity were also predicted using a logistic regression model for bowel V10 suggested by Holyoake et al. [4]. A dose of 10 Gy in 1.8 Gy fractions using LCRT is equivalent to a dose of 8.9 Gy in 5 Gy fractions using SCRT. For all pre- and post-treatment bowels in fx1 and fx5, and for the intrafractional differences in both fx1 and fx5 and for the interfractional differences between fx1 and fx5, this model predicted statistically significantly higher probabilities for grade ≥ 3 acute bowel toxicity than the model provided by Banerjee et al. ($p < 0.01$). For all patients, this model predicted similar probabilities of grade ≥ 3 acute bowel toxicity for the V8.9 pre- and post-treatment bowel loops in both fx1 and fx5. The median (Q1:Q3) probability was 0.12 (0.11:0.18) for the pre-treatment bowel loops in fx1, 0.11 (0.10:0.18) for the post-treatment bowel loops in fx1, 0.16 (0.15:0.23) for the pre-treatment bowel loops in fx5 and 0.15 (0.12:0.21) for the post-treatment bowel loops in fx5. For the intrafractional differences in toxicity risks, the estimate was 0.01 (0.01:0.02) in fx1 and similar 0.02 (0.01:0.02) in fx5. For the interfractional differences in toxicity risks, the estimate was -0.02 (-0.05:0.00) between the pre-treatment bowel loops in fx1 and fx5, and similar -0.02 (-0.04:0.00) between the post-treatment bowel loops in fx1 and fx5. In total, only four pre- and post-treatment bowels in fx1 and fx5 (three patients) had predicted probabilities $< 10\%$ of grade ≥ 3 acute bowel toxicity, and no patients had predicted probabilities $< 5\%$. Two patients had an intrafractional difference in toxicity risk $\geq 5\%$ ($< 10\%$) and one patient had an intrafractional difference $\geq 10\%$. Two patients had an interfractional difference in toxicity risk $\geq 5\%$ ($< 10\%$) for both their pre- and post-treatment bowel loops, where their bowels in fx5 had greater predicted risks of toxicity than their bowels in fx1. Six pre- and post-treatment bowels (four patients) had an interfractional difference in toxicity risk $\geq 10\%$, one with greater predicted toxicity risk in fx1 than fx5.

Chapter 5

Discussion

This study focused on bowel motion during online ART of rectal cancer. In the context of online ART, where the treatment plan is altered in order to account for daily variations in the patient's anatomy, the intrafractional motion is one of the main driver of treatment margins. To our knowledge, not many studies have examined intrafractional bowel motion. Our study quantified intra- and interfractional bowel motion in two RT fractions of 15 rectal cancer patients, and further assessed the impact on dosimetry and toxicity.

Segmentation of anatomical structures in CBCTs are time-consuming, and the poor image quality makes segmenting the bowel loops and distinguishing between the small bowel and colon challenging, resulting in segmentation uncertainties. How long time delineator one spent on segmenting all the structures of interest in a CBCT scan varied depending on the scan length, the amount of air in the bowels causing artifacts, and the delineator's segmentation experience i.e. how many scans delineator one had segmented before. Generally, the time spent on delineating the bladder, small bowel, colon and rectum in one CBCT scan was approximately 5-6 hours. This is far too long to be performed in treatment contexts, especially for online adaptation, and was the motivation to initially investigate an algorithm to increase the segmentation efficiency.

5.1 Segmentations

We used two segmentators to compare manual delineation variations against the performance of TS. When comparing segmentations of the bladder, small bowel, colon and bowel performed by delineator one, two and TS, geometric evaluations showed higher agreements between the manually segmentations than between the segmentations performed by delineator one and TS. The dosimetric evaluation confirmed the reduced performance of TS compared to manual bowel segmentations. One could dis-

cuss whether the segmentations used to calculate agreements between the manual delineators, and thus the measured segmentation uncertainty, are biased, since delineator one was trained by delineator two. During segmentation in fx1, delineator one and two did not seek assistance in each other's segmentations, since this fraction was used to quantify the segmentation uncertainty. However, delineator one and two discussed the CBCTs of some patients when delineator one had doubts about how to interpret the imaged anatomy.

TS v1 was utilised to automatically segment the structures of interest. There are multiple other AI algorithms similar to TS available for auto-segmentation. TS was chosen due to the algorithm being open and available online. The publishers of TS reported that TS v1 achieved a DSC of 0.943 when evaluating its performance on the test set which included different clinical data [101]. TotalSegmentator version two (v2) was released during the late autumn of 2023. V2 contains, among others, retrained models that are trained on improved training datasets and can contour more structures than v1 [99]. The publishers expressed that an open problem present in v2 as for v1 is that the ground truth colon and small bowel segmentations used to train and evaluate the algorithm sometimes were bad due to chaotic and wounding bowels, making it hard to distinguish between the small bowel and colon [99]. As a result, TS also becomes inaccurate in segmenting the small bowel and colon. Also, the model is trained on CTs, and would thus be expected to perform worse in CBCTs of reduced image quality.

5.1.1 Choice of evaluation metrics for segmentations

Multiple different metrics are available for geometric evaluation of segmentation performances [79]. Taha et al. have provided a guide for selecting evaluation metrics based on the objective of the evaluation [79]. Overlap based metrics such as DSC are recommended when outliers are likely to be present in the segmentations or when the absolute volume of the segmentations are important [79]. Distance based metrics such as HD95 are recommended when general alignments of the segmentations and the accuracies of the boundaries are important [79]. In our study, all these factors are of importance, since the aim was to determine the consistency of the different segmentators which includes both accuracy in identifying the structures of interest and their extent including their boundaries.

The DSC is the most commonly utilised metric for validating medical volume segmentations, and has been suggested to be included in any segmentation evaluation [79]. However, DSC is sensitive to the size of the segmentation relative to the background. When segmenting small structures, accurately identifying the structures' exact extent can be challenging compared to greater sized structures. Comparing segmentations,

small errors in delineations of small structures may cause them to erroneously not overlap, and the DSC will indicate poor segmentation performances even if the segmentations lie close to each other. DSC also fails to discover boundary errors. Therefore, distance based metrics such as HD95 should be provided together with the DSC to account for overlap errors and to provide information about the segmentation boundaries and shape similarities [79]. To illustrate the complementary information, Figure 4.7 a) illustrates a slice where only a small part of the bowel was present. TS did not identify an equally large part of the bowel as delineator one did, resulting in a DSC of 0.36, even though the segmentations are located next to each other, partially overlapping. The HD95 value, however, does not indicate an equally bad segmentation performance. Furthermore, Table 4.1 shows that the segmentation comparison resulting in the worst DSC does not mean worst HD95 value. For the colon, the DSC between delineator one and TS resulted in the poorest median value compared to the other colon segmentation comparisons, but not the worst HD95 value, due to the different properties of the metrics, evaluating similarity between structures in various ways. Segmentation performances resulting in DSCs and HD95 values between 0.90-1.00 and 0.00-10.00 mm, respectively, were in our study interpreted as segmentations of good agreements.

TS's segmentation performance was also evaluated dosimetrically to examine if its segmentations were of adequate performance to measure similar dose distributions as the manually segmentations. Poor geometrically agreements between structures does not necessarily mean that their dosimetric estimates differ. As in the example of Figure 4.7 a), the DSC indicated inaccurate bowel segmentation by TS, however, investigating the dose distribution illustrated in Figure 4.7 b) revealed that TS's and delineator one's segmentations in the current slice resulted in an approximately equal dosimetric estimate for the bowel.

A weakness of the in-house Python script used to calculate DSCs and HD95 values between structure segmentations is that it only performs calculations in 2D between slices where both segmentations are present, and a slice with only one segmentation is ignored. When evaluating segmentations in 2D, the DSC in these slices would equal zero, but the HD95 value would be hard to find since the distance between two segmented structures does not exist if only one of the segmentations are present. The script was used since it already existed and was utilised in the clinic to evaluate segmentations. When comparing bowel segmentations by delineator one and TS, the segmented bowels' caudal extension on average (range) deviated by 1 CBCT slice (0-4), and 1 slice (0-2) when comparing the segmentations by delineator one and two. Comparing the manual segmented pre- and post-treatment bowels in fx1 and fx5, and the pre-treatment bowels between fx1 and fx5, the slice differences in caudal extensions were

2 (0-8), 3 (0-9) and 4 (1-12), respectively. Thus, the DSCs and HD95 values measuring intra- and interfractional bowel motion have, to some extent, underestimated motions. If I were to do something differently, I would have created a script to calculate DSCs and HD95 values in 3D, measuring differences in the cranial-caudal direction as well as in the transversal slices.

5.1.2 Segmentation performance and evaluation

The differences in the total segmented bowel volumes for each bowel segmentation in all patients confirmed that TS failed to identify multiple bowel loops and underestimated their volumes. The total segmented bowel volume were not statistically significant different between delineator one and two, but statistically significantly lower for TS's segmentation. This is the volume of TS's post-processed bowels, and its original segmented volume was even lower (Table B6.1).

As well as being time consuming, it was experienced during segmentation that dividing the bowel into the small bowel and colon introduced an increased amount of segmentation uncertainties, even though the DSCs between delineator one's and two's small bowel and colon segmentations were above the threshold of 0.80. Qualitative analyses revealed that TS struggled to distinguish between the small bowel and colon, which was quantitatively confirmed by the DSCs and HD95 values between delineator one's and TS's small bowel and colon segmentations (Table 4.1), where all DSCs were below 0.80. TS's segmentations could, however, be useful as a starting point for the manual segmentations since it delineates some veins, arteries and muscles that might be located close to the bowels. For instance, in some CBCT slices, it could be hard to distinguish the iliac arteries or veins from the small bowel, especially in noisy CBCTs, and TS's segmentations could be used to pinpoint which imaged anatomy was likely to be an artery or a vein and what was likely to be bowel.

The DSCs and HD95 values for the bowel segmentations per post-CBCT slice in fx1, per patient (Figure 4.5 and 4.6 respectively) showed statistically significantly better consistency between the manual segmentations than between delineator one and TS. In the majority of the patients, the IQRs defining the boxes in the segmentation performance plots were also smaller for the manual segmentations, indicating less variations of segmentation agreements in different anatomical regions (caudal vs. cranial slices) compared to between delineator one and TS. The DSCs and HD95 values for the manual bowel segmentations compared well with the segmentation agreements of the manually segmented bladder and rectum, which are organs more visible and less challenging to segment in CBCTs. Also, the results compared well with literature data regarding inter-observer variations of the bowel, such as Perna et al. who compared

manual bowel segmentations in both kVCTs and MVCTs performed by four different observers resulting in mean DSCs of approximately 0.89 and 0.82, respectively [108].

The dosimetric evaluation revealed that the differences between deposited doses in delineator one's and TS's bowel segmentations were statistically significant for low dose levels. For higher dose levels, the differences were smaller or could not be observed at all, and statistically significant differences were not found. These results are expected because the bowels receiving high doses in RT of rectal cancer are normally located caudally within the pelvis, where the bowel loops are more visible and therefore easier to identify and segment both manually and automatically. The bowels located more cranial, receiving lower doses, are often harder to identify, and the cranial CBCT slices were observed to be occupied by more artifacts than the caudal slices (Figure 4.1), resulting in poorer segmentation performances, especially by TS. Low doses are important in predicting acute small bowel toxicities, as shown by, among others, Holyoake et al. [4]. As illustrated in Figure 4.12, TS's bowel segmentations underestimated the bowel volumes exposed to at least 12.6 Gy, which is predictive of grade ≥ 3 acute small bowel toxicity in SCRT, shown for LCRT by Banerjee et al. [92]. Consequently, TS would underestimate the probabilities of grade ≥ 3 acute bowel toxicity, which supports our choice to not further utilise TS's segmentations.

5.1.3 Segmentation performance in different genders

During segmentation, it was observed that TS appeared to be more accurate in segmenting the bladder, small bowel and colon in male patients compared to female patients (see Figure D6.4 comparing DSCs and HD95 values in male and female patients, calculated between the segmentations of delineator one and TS on pre- and post-CBCTs of fx1, and Figure D6.5 for example segmentations in female patients). When segmenting the colon and the bladder in female patients, TS frequently included parts of the cervix-uterus. This seems to have influenced the DSCs and HD95 values for the bladder, resulting in lower and higher measures, respectively, in female compared to male patients. The bladder is known to be a relatively simple structure to identify in CBCTs and the manual delineators rarely had troubles when segmenting bladders. The geometrical comparisons combined with patient specific observations may indicate that TS's segmentations in males are slightly more accurate than in females. We therefore recommend additional caution when using TS for female patients.

TS v1 was trained on a greater amount of CTs of male (≈ 700) than female (≈ 500) patients [101]. Therefore, we expected the DSCs and HD95 values between delineator one's and TS's segmentations to indicate better segmentation performance in male than female patients, as discussed for the bladder. However, the median DSCs and HD95

values do not appear to follow such a trend for all structures. This is explained by the limitation of the script, ignoring slices where only one delineator segmented a structure (see section 5.1.1). As illustrated in Figure D6.5 b), in female patients, caudally colon segmentations performed by TS sometimes confused the cervix-uterus with the colon (rectum as defined by delineator one). This confusion would not contribute to the DSCs or HD95 values of the colon or the bowel since segmentations where delineator one segmented rectum instead of colon were not included in the computations, due to variations in the caudally extents of the colon between the segmentations.

As our study included five female and ten male patients, no statistical analyses were made on gender differences. The newly released TS v2 segments the prostate, but not the cervix-uterus or the ovaries of the female anatomy [99]. TS's segmentations were not utilised to quantify intra- or interfractonal bowel motion or the motion's impact on dosimetry and toxicity, and neither to investigate the bladder volume's impact on bowel exposure. Therefore, TS's inaccuracies in female patients did not further influence the results in our study.

5.2 Intra- and interfractional bowel motion

We used the measures of inter-observer variations between the manually and automatically bowel segmentations to determine whether the manually and/or automatically segmentations were able to measure bowel motion. The DSCs and HD95 values for the manual bowel segmentations representing segmentation uncertainties (median DSC 0.90, HD95 value 7.69), were statistically significantly higher and lower, respectively, than the same metrics measuring intrafractional bowel motion for all patients. This is also illustrated in the differences between the median values in Figure 4.5 and 4.6, suggesting that the manual segmentation uncertainties were lower than the intrafractional bowel motion, and thus managed to measure the bowel motion. The same did not apply to TS's bowel segmentations, which introduced greater uncertainties related to it's segmentations than the uncertainties caused by bowel motion. Therefore, it was decided to carry on with manual segmentations in the subsequent contouring of the CBCTs in fx5, even though TS resulted in bowel DSCs > 0.80, the often used threshold for acceptable segmentation performances. Any bowel motion resulting in a higher DSC or lower HD95 value than the manual segmentation uncertainties can not be measured by the manual segmentations. For instance, the DSCs for patient ER2 measuring intrafractional bowel motion in fx1 and fx5 were higher than the uncertainty measure (Figure 4.13). Therefore, the DSCs can not indicate whether the measured location shifts of the bowel were due to actual intrafractional bowel motion or inaccurate segmentations.

The interfractional bowel motion between fx1 and fx5 was statistically significantly higher than the intrafractional bowel motion within fx1 and fx5, as expected, since daily variations in bowel motions are affected by daily differences in diets, while intrafractional differences are mostly due to motions such as peristalsis and bladder filling. The HD95 values measuring intrafractional bowel motion in fx1 and fx5 were greater than the mean amplitude of the small bowel peristalsis waves of 6.65 $\pm$ 1.15 mm found by Froehlich et al. [8]. They studied the small bowel on MRI images with contrast [8], resulting in less segmentation uncertainties than in our study, which may explain the discrepancies between our results.

To our knowledge, no studies have investigated intrafractional bowel motion in the pelvis (searched PubMed using the key words "intrafraction", "bowel" and "motion"). One study investigating bowel motion during 3D cine-MRI scans and two studies investigating intrafractional motion in the abdomen were however found [109, 110, 111]. Gaudreault et al. quantified intra- and interfractional motion of the small bowel and colon loops during kidney stereotactic ablative radiotherapy (SABR) [111]. They measured intrafractional motion of the small bowel and colon loops in CBCTs by first determining the shortest distance from the ITV to the small bowel or colon in both pre- and post-CBCTs, and then calculating the differences between these distances from the post- and pre-CBCTs [111]. On average, for all 36 patients, the intrafractional small bowel and colon motion were 4.4 mm $\pm$ 5.2 mm and 2.3 mm $\pm$ 2.6 mm respectively [111]. Using a different approach to measure motion, they found less intrafractional bowel motion than our study, however their measures are not directly comparable to our HD95 values. In contrast to our findings, they did not find a significant difference between the intra- and interfractional small bowel and colon motion. Their interfractional small bowel and colon motion, measured between pre-CBCTs and planning CTs, were 4.4 mm $\pm$ 5.0 mm and 5.1 mm $\pm$ 7.1 mm respectively [111], indicating less interfractional motion than the results of our study. Methodological differences between their study and ours and their investigation of structures in the upper abdomen are assumed to influence the results the most. They studied patients without cancer in their bowel, we examined patients with diseased bowels which may affect the bowels' motion [13].

Several studies have quantified interfractional bowel motion using various approaches. Nuyttens et al. (2001) studied the position of small bowel loops in weekly CT scans of four preoperative rectal cancer patients [85]. They measured small bowel motion as the SD of the distances to small bowel from the posterior bones of the pelvis, which was 27 mm in the anterior-posterior direction at 5 cm inferior to the sacral promontory [85]. The HD95 value measuring interfractional bowel motion in our study

is not directly comparable to their findings since it examines the distance between the borders of different bowel loops while Nuyttens et al. studied distances to a fixed bony structure. However, both results are within the same order of magnitude, measuring interfractional bowel motion of approximately 20-30 mm. In contrast to their study, we investigated the small bowel and colon combined in CBCTs. There were methodological differences between their study and ours, where differences in patient positioning, which may provide different conditions for bowel motions [7], are assumed to influence the results the most. Hysing et al. used location probability maps to assess how many times the bowel, investigating the small bowel and colon loops combined, occupied the same position within the pelvis in weekly CT scans of bladder cancer patients. They found that, on average, the total overall volume visited by the bowel at least once was twice as large as the individual average bowel volume found in the CT scans [112]. Additionally, the overall volume visited by the bowel in all repeated scans was on average 20% of the average bowel volume [112]. Sanguineti et al. studied the small bowel and colon combined in weekly CT scans during pelvic IMRT of prostate cancer [90]. On average, for all patients, only 19.2% of the bowel loops were located at the same position at each weekly CT scan compared to the planning CT, consistent with the findings of Hysing et al. [90]. The DSC is not directly comparable to the findings of Hysing et al. and Sanguineti et al. since the DSC calculates the overlap between two bowel segmentations while Hysing et al. and Sanguineti et al. investigated how many times a voxel in the CTs was occupied by bowel and the union of all segmented bowel loops, respectively [90, 112]. Their overlap based methods measured greater interfractional bowel motion than our study was able to detect between CBCTs of two fractions. However, they studied CTs, some with oral contrast, of patients without cancer in their bowels. We investigated CBCTs without oral contrast, which limits the amount of visible bowel loops, of patients with diseased bowels which may influence how the bowels move [13]. It is not surprising that our study seems to have missed some details in the CBCTs, and consequently measured less motion, compared to others using CTs of enhanced image quality. Additionally, some of the CBCT scans in our study have relatively short scan lengths. Increased amounts of the mobile small bowel could perhaps be seen if the scans had greater cranial extents, possibly resulting in larger motions measured. A limitation of their methods compared to ours is the need of multiple CT scans, e.g. Hysing et al. used eight bowel outlines per patient to create a location probability matrix and in some patients they had to use some outlines twice as eight scans were not available [112].

Li et al. studied interfractional motion of the small bowel and the space encapsulating the possible bowel positions, known as the peritoneal space, in weekly CTs of 24

rectal cancer patients [7]. Due to air in the bowels and hence artifacts in the CBCTs, the individual bowel loops in our study could be hard to identify. Therefore, the bowels of some patients were segmented similar (but not necessarily identical) to the peritoneal space (or bowel bag) structure. Li et al. showed that the individual small bowel loops had significantly larger motion than the peritoneal space [7]. Our results may therefore have underestimated both the intra- and interfractional bowel motion, since, in some patients, a larger area than only the bowel loops had to be segmented to ensure that the entire bowel was included in the segmentation. A larger segmented area allows motion in a smaller area to be detected. In addition to neoadjuvant treated patients, Li et al. included adjuvant treated patients in their study, shown to have less mobile bowel loops than neoadjuvant patients [85]. Differences in patient positioning between their study and ours and their inclusion of both neoadjuvant and adjuvant patients are assumed to influence the bowel motion results.

5.3 Dosimetric consequences of motion

Using adaptive treatments, the intrafractional bowel motion was shown to influence the dose distribution within the bowel.

5.3.1 Intra- and interfractional bowel motion and exposure

Our study investigated the impact of intra- and interfractional bowel motion for adaptive treatments using 12 fields IMRT. For the majority of the patients, the intrafractional bowel motion caused dosimetric benefits for bowels exposed to low doses, with less impacts for bowels exposed to higher doses (Figure 4.18). As earlier explained, the differences in the dose distributions between the low and high doses may be due to segmentation uncertainties in the cranial CBCT slices where the bowels exposed to low doses are located. Additionally, in the low dose regions, cranial bowel segmentations mostly consisted of the small bowel, which is more mobile than the colon. Therefore, the cranial bowels were expected to experience greater intrafractional variations in irradiated volumes. For the high dose regions, caudal loops were more visible in the CBCTs and thus could be more accurately segmented. Also, since the transition between the rectum and the colon (sigmoid) is continuous, some parts of the colon must lie close to the rectum both pre- and post-treatment and thus receive high doses both pre- and post-treatment in all fractions. Then, the intrafractional difference in irradiated bowel volume will be small, which may partly explain the low volume differences among the highest dose levels.

There were considerable variations in both intra- and interfractional bowel motion

in the investigated cohort of patients (figs. 4.13 to 4.15). Consequently, the intra- and interfractional dosimetric consequences of motion varied both between patients and for individual patients between fx1 and fx5 (Figure 4.18). Seven individual patients had a fixed bowel loop permanently located close to the CTV. An example is patient ER6, illustrated in Figure 4.24. In fx5, the DVH in Figure 4.18 for ER6's dose-volume-differences centers around zero for doses > 10 Gy, because of very little intrafractional motion beyond these dose levels. In fx1, the pre- and post-treatment bowel's DVHs had approximately similar shape (Figure 4.22), but a greater volume of the pre-treatment bowel received dose, for all dose levels, compared to the post-treatment bowel, which explains why the dose-volume-difference curve in Figure 4.18 a) flattens out at a volume difference above zero. The fixed bowel loop received high doses in all fractions. Therefore, it is reasonable that the DVHs for dose-volume-differences illustrate small volume differences for high dose levels.

Several studies have found relations between interfractional bowel motion and dose distribution [7, 113, 114]. Kvinnland et al., Li et al. and Nuytttens et al. (2004) found greater interfractional variations in irradiated bowel volumes, even for high dose levels. Nuyttens et al. (2004) found that the SD of the small bowel volume exposed to $\geq 95\%$ of the prescribed dose ranged from 2 to 67 cc using IMRT [114]. Kvinnsland et al. studied the uncertainties in bowel DVHs caused by interfractional bowel motion in urinary bladder cancer patients treated with 3DCRT [113]. For the dose-volume parameters V48%, V77% and V84%, they found mean values of irradiated bowel volumes ranging from 77-336 cc, 52-250 cc and 38-243 cc respectively [113]. Intrafractionally, our study did not observe an equally large variation in bowel exposure, especially for high dose levels, as illustrated in Figure 4.19 where only outliers reached equally high volume differences as their results. Different RT techniques and the use of CT compared to our CBCTs are factors affecting our results. Similar to a study by Li et al. who investigated the impact of interfractional small bowel motion on small bowel dose exposure, we found that the intrafractional variations in bowel exposure were greatest for the smallest dose levels, decreasing with increasing dose [7]. Ignoring some outliers, their results are comparable to the variations between our patients' DVHs for high dose levels, where the intrafractional volume differences lie approximately between -10 and 20 cc for V20 and V22.5, illustrated in Figure 4.19 b).

5.3.2 Impact of bladder filling on bowel exposure

Rectal cancer patients may experience variations in bladder filling during online ART which can contribute to the bowels' intrafractional changes in position. In our study, great variations in bladder volumes and intrafractional bladder fillings were indeed

found both between patients and between fractions for individual patients, possibly explained by the absence of hydration instructions. Smith et al. studied intrafractional bladder filling in ten prostate cancer patients treated on an MR-linac using bladder filling instructions [115]. Session and verification MRIs were acquired pre- and post-optimisation of the adaptive plans, resulting in a median (range) bladder volume of 121 cc (46-708 cc) and 211 cc (70-933 cc) respectively [115]. Hence, they measured an intrafractional median difference in bladder volume of 90 cc, which is greater than our median intrafractional differences of -2.9 cc in fx1 and -0.4 cc in fx5 (see bladder volumes in Table C6.2). This difference may be caused by their increased time between the image acquisitions (median (range) 25 min (19-34 min)) compared to ours, and their use of a hydration regimen.

Nuyttens et al. (2001) argued that reducing the small bowel volume in the anterior midpelvis may have an impact on small bowel toxicity [85], which is dependent on small bowel V12.6 according to Banerjee et al. [92] using 5×5 Gy fractionation. Therefore, and because an increased bladder volume tends to lift the bowel in the cranial direction, away from the RT target [85, 86], we expected a greater bladder volume to reduce bowel exposure. No statistically significant relationships between pre-treatment bladder volume and bowel exposure were however found (see section 4.3.2). Pre-treatment bladders and bowels were investigated, instead of post-treatment structures, due to the possibility to modify the bladder volumes prior to treatment delivery if a relation were to be found. The linear regression analyses resulted in a slope of -0.19 and 0.01 when investigating the linear relationships between the pre-treatment bladder volumes and pre-treatment bowel V10 and V20, respectively (Table 4.2). Therefore, the linear regression analyses suggested that a greater bladder volume reduced the bowel volume exposed to at least 10.0 Gy, and slightly increased the exposure for bowel V20. However, the p-values for the Pearson correlation coefficients indicated no statistically significant linear relationships. Hence, no trends between pre-treatment bladder volume and bowel exposure were found. Perhaps a different relation could be found between the bladder volumes and irradiated bowel volumes if our study included an increased amount of patients with full bladders. Four patients in fx1 and only one patient in fx5 had pre-treatment bladder volumes ≥ 300 cc. Nuyttens et al. (2004) investigated the impact of bladder filling on the volume of irradiated small bowel in rectal cancer patients treated with full bladder instructions [114]. Using IMRT, they showed that a full bladder decreased the small bowel volume inside the 90% isodose line by 190% in four preoperative patients [114]. Thus, Nuyttens et al. measured a relation between bladder volume and small bowel exposure for high doses that we were not able to find. Methodological differences in RT techniques and bowel definitions between

their study and our are assumed to influence the results. Perhaps we lost some information regarding the impact of bladder volume on bowel exposure when we combined the small bowel and colon.

The linear regression analyses relating intrafractional differences in bladder volumes to the intrafractional differences in bowel volumes exposed to at least 10.0 Gy and 20.0 Gy suggested a slight relation between an intrafractional increase in bladder volume and an intrafractional decrease in bowel exposure (Table 4.2, Figure 4.27). Comparing the slopes from the linear regression analyses indicates that an intrafractional change in bladder filling has greater impact on an intrafractional change in bowel exposure compared to the linear relation between pre-treatment bladder and irradiated bowel volumes. However, a statistically significant correlation was only found between the intrafractional changes in bladder volumes and bowel V20s ($r = -0.38$, $p < 0.05$). Hence, an intrafractional increase in bladder volume tends to result in an intrafractional decrease in bowel exposure for high dose levels. Similar to our results for bowel V10 (V40%), Xu et al. did not find statistically significant relations between interfractional changes in bladder volumes and changes in small bowel loop V30% or bowel bag V89% during IMRT of gynecologic cancer patients using linear regression [91]. However, they studied patients with prior hysterectomies (surgically removal of the uterus) [91], which can affect the bowel motion [116]. Their investigation of weekly CTs and the different patient cohort and bowel definition in their study compared to our are assumed to influence the results.

Some studies have also investigated the relation between bladder volume and small bowel position in CTs, without dosimetric evaluations [85, 86]. I chose to highlight their findings because the position of the small bowel indicates the potential dose exposed volume. Nuyttens et al. (2001) showed that the most inferior small bowel position was correlated with the bladder volume, and a full bladder reduced the small bowel volume in the anterior midpelvis by 140% [85]. Zhang et al. found that the small bowel volume within a limited region of interest of the pelvis decreased as the bladder volume increased in cervical cancer patients [86], a similar relation as observed by Nuyttens et al. Qualitative analyses of individual patients in our study agreed with their findings where a reduced bladder volume allowed a greater amount of the bowel to be located more caudally. A statistically significant relation was, however, not found for the pre-treatment structures, possibly due to issues with small bowel or small bowel and colon combined and few patients.

5.4 Predicting grade ≥ 3 acute toxicity in bowel loops

Multiple studies have shown that RT induced acute small bowel toxicity is related to the volume of irradiated small bowel [92, 95, 96]. Low doses are important in predicting risks of acute small bowel toxicity, as shown by Banerjee et al. in a small, homogeneous group of rectal cancer patients and in a meta-analysis by Holyoake et al. [4, 92]. Logistic regression models for toxicity probability estimation from both these studies were utilised in our study to predict the probabilities of grade ≥ 3 acute toxicity for the manually segmented bowel loops. We primarily focused on the model generated by Banerjee et al., who provided a logistic regression model for acute toxicity probability estimation for individual bowel loops [92]. The model by Holyoake et al. was used as a comparison. Banerjee et al. studied rectal cancer patients, all treated similarly with curative neoadjuvant CRT [92]. Disregarding the females with prior hysterectomy, the 3DCRT treatment technique, and the use of a bellyboard and prone patient positioning for some patients (which can affect the position and motion of the bowel), the patient cohort studied by Banerjee et al. is more similar to the patient group in our study compared to the patients included in the analysis by Holyoake et al. The meta-analysis by Holyoake et al. is a large study, analysing previous published data of small bowel dose-volumes and toxicity outcomes, including the findings by Banerjee et al. [4]. A benefit of the model provided by Holyoake et al. is therefore that it relates toxicity to a large patient group, using a wide range of data reporting DVH parameters and toxicity outcomes for patients treated with conventionally fractionated RT [4]. However, the meta-analysis, introducing multiple uncertainties due to inconsistency in their studied data, which impacts their reported dose-volume and toxicity relations, might be less suitable to predict toxicity probabilities for the patients included in our study. Our toxicity probabilities estimated using the model by Banerjee et al. probably resulted in more realistic toxicity estimates compared to the corresponding estimates using the model by Holyoake et al., however these results are also highly affected by methodological uncertainties. The extrapolations in the fractionation schemes from LCRT to SCRT are assumed to result in the greatest uncertainties in our toxicity risk estimates. Additionally, both models are generated for the small bowel, we studied the small bowel and colon combined. For LCCRT of rectal cancer, Gunnlaugsson et al. studied the relationship between irradiated small bowel and colon volumes and enteritis (inflammation), an acute RT induced side effect [117]. They showed that the irradiated small bowel volume was statistically significantly associated with CRT induced enteritis [117]. They did not find a statistically significant relation between the irradiated colon or peritoneal space volume and acute enteritis [117]. However, patients with diarrhoea, studied as a

measure of enteritis, of grade ≥ 2 had greater volumes of irradiated colon or peritoneal space for all dose levels compared to patients with diarrhoea of lower grades [117]. Our bowel structure consisted of a greater volume than the volume of the small bowel alone. If the colon is as radiosensitive as the small bowel, our toxicity estimates remain approximately unaffected by not distinguishing between small bowel and colon. However, if the colon is less radiosensitive, the models can overestimate the toxicity risks since the models' input will represent a greater volume of irradiated radiosensitive bowels than the actual volume of radiosensitive bowels present. On the other hand, if the colon is more radiosensitive than the small bowel, the toxicity estimates will probably be underestimated.

When comparing different treatment arms, a difference of $\geq 5\%$ in grade ≥ 3 toxicity outcomes between the treatment arms is commonly chosen to indicate a clinically significant difference. Using the model by Banerjee et al., the majority of the patients had low risks for grade ≥ 3 acute bowel toxicity, i.e. risks below 5% (Figure 4.29). We estimated toxicity risks at specific moments in time; we investigated the toxicity risk at the moments when the bowel was positioned in a certain way during the pre-CBCT scans, and similarly at the moments during the post-CBCT scans. An intra- or interfractional change in the toxicity estimate occurs somewhere between these CBCT acquisitions. Hence, the actual toxicity estimate for a patient in one fraction has a value between the estimated toxicity risks for the pre- and post-treatment bowels. However, the overall toxicity risk estimate for grade ≥ 3 acute bowel toxicity for a patient is a combination of the toxicity risk estimates from all five RT fractions, we only estimated the risks for two of these fractions.

In order to compare with a study by Li et al. of interfractional changes in toxicity estimates [7], I used their reported small bowel V15 to predict probabilities of small bowel toxicity using the model by Banerjee et al. This resulted in small bowel toxicity estimates of 5.8% and 2.3% from their CTs acquired one and two weeks after treatment planning, respectively. These results are similar to, but slightly higher, than our toxicity risk estimates. Further, the toxicity estimates based on the two investigated CTs from the study by Li et al. resulted in an interfractional difference of 3.5%. These CTs from the study by Li et al. were approximately separated by the same time as our CBCTs from fx1 and fx5. Their interfractional difference is greater than our measured both intra- and interfractional differences in toxicity risk estimates when evaluating all patients combined. Only one patient in our study had an intra- and interfractional difference in toxicity risk estimates equally high as the interfractional difference from the study by Li et al. Similarities between their study and our are reasonable since their variations in dose-volume parameters were similar to ours, as earlier discussed (Section

5.3.1). However, different treatment techniques and issues due to our combined small bowel and colon, as explained earlier, may cause the slight discrepancies their results and ours.

In patients with intra- and interfractional decreases in toxicity risk estimates, bowel motion caused the bowel to be in a more favorable location relative to the dose distribution within the pelvis, reducing the volumes exposed to doses predictive of toxicity (example ER13). However, in CBCTs occupied by artifacts, either in the entire scan or in the cranial slices, the manual bowel segmentations were inaccurate. Therefore, individual bowel loops were not identified, and the bowels were segmented similar to bowel bag structures (example Figure 4.2 for ER13). When either the pre- or post-CBCTs within one fraction, or between fractions, were more degraded by artifacts than the other, inaccurate segmentations in the CBCTs of worst image quality most likely overestimated the bowel volumes exposed to all dose levels compared to the segmentations in the CBCTs of better image quality. It is reasonable that a segmentation similar to a bowel bag structure, increasing the total segmented volume, results in inaccurate toxicity estimates when using a model designed for individual bowel loops. Banerjee et al. also provided a logistic regression curve for toxicity probability estimation using peritoneal space delineation [92]. Perhaps the logistic regression curve for peritoneal space V15 could predict more accurate toxicity probabilities for the patients with the CBCTs of worst image quality, such as patient ER13.

There are multiple patients with greater both intra- and interfractional bowel motion who does not have clinically significant intra- or interfractional differences in toxicity risk estimates. Intra- or interfractional location shifts of the bowel that does not move the bowel in to or out of pelvic regions exposed to $\geq$ 12.6 Gy will not impact the toxicity risk estimates using the chosen model. An example is patient ER14, where the toxicity risk estimate did not change intrafractionally in fx5 (Figure 4.30 b) despite the presence of intrafractional bowel motion (Figure 4.13 and 4.14).

Seven patients had a fixed bowel loop located close to the CTV in all fractions. When the same bowel loop is permanently exposed to high doses, the probability of radiation induced damages increases. The model used to estimate toxicity risks does not consider if the same bowel segment always is exposed to high doses. Nor does it consider if the volume it predicts on is small bowel or colon. Certain of the patients with the fixed bowel loops resulted in some of the highest observed toxicity risk estimates (Figure 4.30). This is caused by a greater total bowel volume exposed to doses predictable of toxicity, not explicitly due to the fixed bowel loop close to the CTV. However, some of the other patients with the fixed bowel loops did not result in high toxicity risk estimates, due to a small overall volume of bowels exposed to doses predictive of

toxicity. These patients may not be at risk of developing grade ≥ 3 acute bowel toxicity. Gaudreault et al. observed that for patients suffering from bowel stricture, which is an RT induced side effect, the colon was inside the planned 100% isodose in all pre- and post-CBCTs using non-adaptive treatments [111]. In the case of an eventual dose escalation, patients with a fixed bowel loop close to the CTV may be exposed to severe side effects such as perforation, which can occur when the same bowel segment is repeatedly exposed to high doses [2]. With the current strategy, there are no indications of a risk of bowel perforation.

Banerjee et al. did not account for daily variations in the patients' anatomy and the intra- and interfractional bowel motion found in our study due to their investigations of planning CTs only [92]. Sanguineti et al. showed, during weekly CT scans, that approximately 280 cc of the bowel loops were not located at the same place during treatment as in the planning CTs [90]. Therefore, the small bowel DVHs found by Banerjee et al. are not necessarily representative of the actual irradiated bowel volumes during treatment, for all dose levels. For instance, the high doses might have been distributed over several different bowel segments during treatment, requiring repeated CT scans during the treatment course to be detected. Our study investigated daily CBCTs where intra- and interfractional bowel motion are accounted for when examining the dose distributions. Consequently, the DVH parameters in our study are likely better representatives of the dose distributions within the bowels during treatment than the DVH parameters found by Banerjee et al., though with a larger delineation uncertainty due to the image quality difference in CBCTs and CTs. While none of the patients included in our study had previously been operated in the pelvis, some of the females included in the study by Banerjee et al. had undergone prior hysterectomies, which typically allows more bowel to be located within the radiation field during rectal cancer RT [92]. A greater volume of these patients' bowel volumes might have been exposed to high doses compared to the bowel volumes of the patients in our study. Also, hysterectomy can affect the bowel function in some patients, which in turn can affect the motion of the bowel [116].

Probabilities of grade ≥ 3 acute bowel toxicity were also predicted using the logistic regression model provided by Holyoake et al. [4], resulting in statistically significantly higher toxicity probabilities than the results from the model by Banerjee et al., for both the pre- and post-treatment bowels in both fx1 and fx5. There are multiple uncertainties related to both models for toxicity risk estimation, which challenges the reliability of our toxicity risk estimates. Both models were generated based on LCRT, while our treatments were delivered as SCRT. Using the model provided by Holyoake et al., an approximation were performed to determine the dose equivalent to 10 Gy using 5×5

Gy fractionation. In the calculation, it was assumed that Holyoake et al. studied LCRT of 28×1.8 Gy, however some of the publications they analysed reported different treatment schemes. Our bowels were not limited to the individual bowel loops only because the poor image quality of the CBCTs prevented us from accurately identifying each loop. Thus, the predicted toxicity probabilities in our study are affected by substantial uncertainties in the segmented volumes. There are also methodological uncertainties related to the use of WebPlotDigitizer in extracting the numerical data for the logistic regression curve generated by Banerjee et al. (Figure 3.5), however this uncertainty is small compared to the segmentation uncertainty.

5.5 Study limitations and benefits

Our study benefits from detailed volumetric images for all treatment fractions. Hence, we were able to analyse both intrafractional bowel motion and dose distributions within the bowels during these fractions. Our study provides data for dose-response modelling for 5×5 Gy fractionation. Since the patients' actual toxicity outcomes were not available, and since some of our patients were treated with LCRT (see Table 3.1), we were not able to generate dose-response models for SCRT. In the future, we could use a dose-response model for the bowel generated for SCRT. Then, our toxicity risk estimates would not be influenced by the uncertainties due to extrapolating the LCRT fractionation and the dose-response models.

The small number of patients, and the total number of 60 segmented CBCT scans of poor image quality, are the main limitations to our study. Only two fractions per patient were examined due to the time spent on delineating the structures of interest. Fx1 and fx5 were chosen because they were separated with the longest time interval. Increasing the number of patients and the investigated number of fractions per patient could have enabled us to identify more trends in the data and increase the strengths of the statistical analyses.

Artifacts in the CBCTs, mainly observed as streak artifacts due to bowel gas, reduced the accuracy of both manually and automatically segmentations, consequently affecting the measured intra- and interfractional bowel motions, the dose distributions and the predicted toxicity risks. Large volumes were segmented to ensure inclusion of the entire bowel due to the challenges of delineating individual bowel loops in the poor quality CBCTs. Consequently, the measured bowel volume within the pelvis are inaccurate and the precision of the motion measures are reduced. Moreover, the segmentation of a larger bowel volume may indicate that a higher volume receives dose than the actual volume receiving dose. Recently, HyperSight™imaging became avail-

able at HUH. HyperSight™offers a reduced acquisition time compared to the conventional linac-based CBCT and thus reduces motion-related artifacts and patient discomfort [118]. It also offers a larger FOV and advanced reconstruction algorithms to increase image quality and reduce artifacts, improving spatial and contrast resolution. Additionally, HyperSight™acquires images with accurate HUs which directly can be used in calculating dose distributions during treatment planning [118]. Consequently, HyperSight™can be useful in online ART to adjust for anatomical changes at each treatment fraction [119]. If images acquired using HyperSight™instead of the conventional linac-based CBCT were available in our study, our results would be less affected by uncertainties. Additionally, segmenting the bowel would probably be less time-consuming. Following, perhaps multiple fractions per patient could be included in our study, and/or we could potentially distinguish between the small bowel and colon and study these structures separate.

Some patients, such as patient ER10, had more artifacts present in one of the CBCT scans compared to the other within one fraction, as illustrated in Figure 4.21. Inaccurate delineations where a notably larger bowel volume is segmented in the pre- than the post-CBCTs, or vice versa, may result in an overestimation of the intra- and/or interfractional bowel motion. It will especially affect the overlap based metric, since a full overlap between two structures are impossible when they are of unequal sizes. Three and five patients in fx1 and fx5 respectively had distinct unequal amounts of artifacts in their pre- and post-CBCTs, affecting the total segmented bowel volumes. Additionally, when copying structures from the post-CBCTs to the pre-CBCTs or from the CBCTs of fx5 to fx1, some segmentations, mostly in the cranial slices, were distorted due to differences in treatment couch pitch between the CBCT scans, affecting the interpolated segmented volume.

The chosen method to quantify intra- and interfractional bowel motion between pre- and post-CBCTs in two fractions is not an optimal method to measure motion, but it gives an estimate of the bowels' displacements. Four dimensional (4D) MRI with a temporal resolution great enough to record rapid bowel motions, such as the small bowel peristalsis waves occurring approximately between eight and eleven times per minute according to Froehlich et al. [8], could perhaps provide a better method to measure bowel motion. Johansson et al. showed that a 4D MRI technique could visualize periodic GI motions in the abdomen, mainly focusing on the stomach, and suggested that a similar method could be used to explore motion in other structures as well [120]. Nevertheless, this approach would require more visits on a diagnostic MR or access to an MR-linac to monitor potential changes of intrafractional motion over a treatment course.

A methodological uncertainty in our study is the technique used to segment the bowels in fx1 and fx5. When delineator one segmented the CBCTs of fx1, the small bowel and colon loops were first segmented separately before they were combined to a common bowel structure. Delineator two divided its previously segmented combined bowel structure into small bowel and colon. During the division, delineator two had to remove some "help lines" that were included in the combined bowel structure to get a realistic structure, that did not make sense to include in the small bowel or colon structure when these were segmented separately. This removal thus led to a reduction in the total segmented volume. In fx5, it was not attempted to distinguish between the small bowel and colon loops, and the combined bowel structure was segmented directly. Therefore, there might be a greater segmented bowel volume in fx5 than fx1, which could have affected the measured intrafractional bowel motion in fx1 and fx5, or the interfractional bowel motion between them. However, the difference between the intrafractional bowel motion in fx1 and fx5 was not statistically significant.

The contouring guidelines provided by RTOG encourages to use oral contrast to distinguish between small bowel and colon in CTs, and did not give any guidelines in how to distinguish small bowel from colon in the absence of oral contrast [77]. Oral contrast was not utilised in this study. In the absence of small bowel contrast, the guidelines suggest delineating the bowel bag, a simpler and faster option instead of delineating the small bowel and colon loops [77]. Since online ART adapts the treatment plan to the current anatomy and therefore the current bowel position, the need for a larger volume that considers all possible bowel locations is reduced. Therefore, an effort was made to segment the individual loops and not the bowel bag, and knowledge about the small bowel and colon anatomy was used to distinguish between them in fx1. In fx5, it was decided not to distinguish between the small bowel and colon, and the bowel was segmented as one combined structure, due to the time spent on delineation in fx1. By distinguishing between the small bowel and colon, more information about bowel motion could possibly be gained, but it would be heavily affected by segmentation uncertainties due to the poor image quality of the available CBCTs. When delineator one had gained bowel segmentation experience, segmenting only the bowel and the bladder in one CBCT scan took about 3-4 hours.

Chapter 6

Conclusions

This thesis presented new data on intrafractional bowel motion relevant for online adaptations. An AI algorithm for auto-segmentation of anatomical structures in CBCT images did not show adequate performance in bowel segmentations, even when post-processed. Therefore, the auto-segmentation algorithm was not able to measure intra- and interfractional bowel motion in CBCTs. Manual segmentations were able to measure both intra- and interfractional bowel motion when the small bowel and colon were segmented as one structure. The small bowel and colon were combined due to the poor image quality of the CBCTs and to reduce the workload. The intrafractional bowel motions in both fx1 and fx5 were, for all patients, significantly lower than the interfractional bowel motion between fx1 and fx5, as expected.

For the majority of the investigated patients and fractions, the intrafractional bowel motion seemed to affect the dose distribution in the bowel during ART of rectal cancer, moving the bowel out of the pelvic regions exposed to doses, especially for low dose levels. There seemed to be a relation between increased bladder volume and reduced bowel V10, but a statistically significant relation was not found. Additionally, a significant relation between bladder volume and irradiated bowel volume was not found for the higher dose level (V20). In contrast, a statistically significant correlation between an intrafractional increase in bladder volume and decrease in bowel exposure was found for volumes exposed to at least 20.0 Gy. However, considering the segmentation uncertainties due to the poor image quality and artifacts present in the CBCTs, and the small number of patients, these results need further investigation, ideally using images of enhanced quality and a larger number of patients and fractions.

Based on the dose-volume parameters considering intrafractional bowel motion, the number of patients with risk estimates $\geq 5\%$ for grade ≥ 3 acute bowel toxicity varied depending on the dose-response model. Because of uncertain dose-response models used for toxicity risk estimation, and uncertainties in the bowel volumes and dose distri-

butions due to segmentation uncertainties, the influence of intrafractional bowel motion on toxicity risk estimates and toxicity outcomes requires further investigation, ideally with a bowel dose-response model for SCRT.

Bibliography

[1] Kavanagh BD, Pan CC, Dawson LA, Das SK, Li XA, Haken RKT, et al. Radiation Dose-Volume Effects in the Stomach and Small Bowel. International Journal of Radiation Oncology Biology Physics. 2010 3;76.

[2] Letschert JJ. The Prevention of Radiation-induced Small Bowel Complications; 1995.

[3] Bahadoer RR, Dijkstra EA, van Etten B, Marijnen CAM, Putter H, Kranenbarg EMK, et al. Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. The Lancet Oncology. 2021 1;22:29-42.

[4] Holyoake DLP, Partridge M, Hawkins MA. Systematic review and meta-analysis of small bowel dose–volume and acute toxicity in conventionally-fractionated rectal cancer radiotherapy. Radiotherapy and Oncology. 2019 9;138:38-44.

[5] Varian Medical Systems. Ethos™Therapy: Intelligent Adaptation Comes to Clinic [Online]. Varian Medical Systems; 2020 Jan 28. [cited 2024 Mar 16]. Available from: https://www.varian.com/ethostm-therapy-intelligent-adaptation-comes-clinic.

[6] Green OL, Henke LE, Hugo GD. Practical Clinical Workflows for Online and Offline Adaptive Radiation Therapy. Seminars in Radiation Oncology. 2019 7;29:219-27.

[7] Li S, Gong Y, Yang Y, Guo Q, Qian J, Tian Y. Evaluation of small bowel motion and feasibility of using the peritoneal space to replace bowel loops for dose constraints during intensity-modulated radiotherapy for rectal cancer. Radiation Oncology. 2020 9;15.

[8] Froehlich JM, Patak MA, Weymarn CV, Juli CF, Zollikofer CL, Wentz KU.

Small bowel motility assessment with magnetic resonance imaging. Journal of Magnetic Resonance Imaging. 2005 4;21:370-5.

[9] Joiner M, van der Kogel A. Basic Clinical Radiobiology. 4th ed. Boca Raton: Taylor Francis Group; 2009.

[10] Harrison K, Pullen H, Welsh C, Oktay O, Alvarez-Valle J, Jena R. Machine Learning for Auto-Segmentation in Radiotherapy Planning. Clinical Oncology. 2022 2;34:74-88.

[11] Helseinformatikk N. Kreft [Online]. Norsk Helseinformatikk;. [updated 2021 Feb 25; cited 2023 Nov 07]. Available from: https://nhi.no/kroppen-var/sykdomsprosesser/kreft/.

[12] Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rödel C, Cervantes A, et al. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2017;28:iv22-40.

[13] PDQ® Adult Treatment Editorial Board. PDQ Rectal Cancer Treatment [Online]. Bethesda, MD: National Cancer Institute;. [updated 2024 Feb 15; cited 2024 Mar 03]. Available from: https://www.cancer.gov/types/colorectal/hp/rectal-treatment-pdq.

[14] Mustafa SMT, Malik V. Current paradigms in rectal cancer. Current Medicine Research and Practice. 2017;7(2):59-66. Available from: https://www.sciencedirect.com/science/article/pii/S2352081717300168.

[15] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2021 5;71:209-49.

[16] Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. International Journal of Cancer. 2021 8;149:778-89.

[17] Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, et al.. Global Cancer Observatory: Cancer Today. Lyon: International Agency for Research on Cancer; 2020. [cited 2023 Oct 16]. Available from: https://gco.iarc.fr/today.

[18] Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors. Termedia Publishing House Ltd.; 2019.

[19] Årsrapport 2022 med resultater og forbedringstiltak fra Nasjonalt kvalitetsregister for tykk- og endetarmskreft. Oslo: Kreftregisteret; 2023. ISBN: 978-82-473-0124-1.

[20] Kreftforeningen. Tarmscreeningsprogrammet - screening for å oppdage tarmkreft [Online]. Kreftforeningen;. [updated 2023 Oct 02; cited 2023 Nov 07]. Available from: https://kreftforeningen.no/forebygging/screening-og-masseundersokelser/tarmscreeningprogrammet/.

[21] National Cancer Institute. TNM staging system [Online];. [cited 2023 Oct 20]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/tnm-staging-system.

[22] Matalon SA, Mamon HJ, Fuchs CS, Doyle LA, Tirumani SH, Ramaiya NH, et al. Anorectal cancer: Critical anatomic and staging distinctions that affect use of radiation therapy. Radiographics. 2015 11;35:2090-107.

[23] Holck P. bukhinnen [Online];. [updated 2023 Jul 11; cited 2023 Oct 30]. Store medisinske leksikon, snl.no. Available from: https://sml.snl.no/bukhinnen.

[24] Hope TA, Gollub MJ, Arya S, Bates DDB, Ganeshan D, Harisinghani M, et al. Rectal cancer lexicon: consensus statement from the society of abdominal radiology rectal anal cancer disease-focused panel. Abdominal Radiology. 2019 11;44:3508-17.

[25] Lambregts D, van Loenhout R, Zijta F, Lahaye M, Beets-Tan R, Smithuis R. Rectal Cancer MR staging 3.0 [Online]. Netherlands: Radiology Departement of the Netherlands Cancer Institute in Amsterdam, the Medical Centre Haaglanden in the Hague and the Alrijne Hospital in Leiderdorp; 2021 Sep 01. [cited 2023 Oct 27]. Available from: https://radiologyassistant.nl/abdomen/rectum/rectal-cancer-mr-staging-1.

[26] Helsedirektoratet. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av kreft i tykktarm og endetarm [nettdokument]. Oslo: Helsedirektoratet; 2017. [updated 2024 Jan 31; cited 2024 Mar 04]. Available from: https://www.helsedirektoratet.no/retningslinjer/kreft-i-tykktarm-og-endetarm-handlingsprogram.

[27] Kreftforeningen. Pakkeforløp [Online]. Kreftforeningen;. [updated 2023 Aug 31; cited 2023 Nov 07]. Available from: https://kreftforeningen.no/rad-og-rettigheter/dine-rettigheter-som-pasient/ventetider-og-pakkeforlop/.

[28] PDQ® Adult Treatment Editorial Board. PDQ Rectal Cancer Treatment [Online]. Bethesda, MD: National Cancer Institute;. [Updated 2023 May 15; cited 2023 Oct 25]. Available from: https://www.cancer.gov/types/colorectal/patient/rectal-treatment-pdq.

[29] Kreftforeningen. Tarmkreft [Online]. Kreftforeningen;. [updated 2023 Nov 02; cited 2023 Nov 09]. Available from: https://kreftforeningen.no/om-kreft/kreftformer/tarmkreft/.

[30] Helsedirektoratet. Pakkeforløp for tykk- og endetarmskreft [Online]. Oslo: Helsedirektoratet; 2014. [updated 2022 Apr 29; cited 2023 Oct 23]. Available from: https://www.helsedirektoratet.no/nasjonale-forlop/tykk-og-endetarmskreft.

[31] Klepp O, Olsen TK. kreftbehandling [Online]; April 14, 2023. [updated 2023 Apr 14; cited 2023 Oct 24]. Store medisinske leksikon, snl.no. Available from: https://sml.snl.no/kreftbehandling.

[32] Delibegovic S. Introduction to Total Mesorectal Excision. Medical archives (Sarajevo, Bosnia and Herzegovina). 2017 12;71:434-8.

[33] Schlichting E. radikaloperasjon [Online]; 2018 Sep 04. [cited 2023 Oct 23]. Store medisinske leksikon, snl.no. Available from: https://sml.snl.no/radikaloperasjon.

[34] Statsforvaltaren i Vestland. Fleire pasientar fekk ikkje forsvarleg helsehjelp gjennom deltaking i Norwait-studien [Online]. Statsforvaltaren i Vestland; 2023 Mar 28. [cited 2024 Mar 27]. Available from: https://www.statsforvalteren.no/vestland/helse-omsorg-og-sosialtenester/helsetenester/fleire-pasientar-fekk-ikkje-forsvarleg-helsehjelp-gjennom-deltaking-i-norwait-studien/.

[35] Dijk THV, Tamas K, Beukema JC, Beets GL, Gelderblom AJ, de Jong KP, et al. Evaluation of short-course radiotherapy followed by neoadjuvant bevacizumab, capecitabine, and oxaliplatin and subsequent radical surgical treatment in primary stage IV rectal cancer. Annals of Oncology. 2013 7;24:1762-9.

[36] Rosen H, Sebesta CG, Sebesta C. Management of Low Anterior Resection Syndrome (LARS) Following Resection for Rectal Cancer. MDPI; 2023.

[37] Peeters KCMJ, van de Velde CJH, Leer JWH, Martijn H, Junggeburt JMC, Kranenbarg EK, et al. Late side effects of short-course preoperative radiotherapy

combined with total mesorectal excision for rectal cancer: Increased bowel dysfunction in irradiated patients - A Dutch Colorectal Cancer Group Study. Journal of Clinical Oncology. 2005;23:6199-206.

[38] Glimelius B, Grönberg H, Järhult J, Wallgren A, Cavallin-Ståhl E. A systematic overview of radiation therapy effects in rectal cancer; 2003.

[39] Sauer R, Becker H, Hohenberger W, Rödel C, Wittekind C, Fietkau R, et al.. Preoperative versus Postoperative Chemoradiotherapy for Rectal Cancer; 2004. Available from: www.nejm.org.

[40] Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: Results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. Journal of Clinical Oncology. 2012 6;30:1926-33.

[41] Tveit KM, Guldvog I, Hagen S, Trondsen E, Harbitz T, Nygaard K, et al. Randomized controlled trial of postoperative radiotherapy and short-term time-scheduled 5-fluorouracil against surgery alone in the treatment of Dukes B and C rectal cancer. Norwegian Adjuvant Rectal Cancer Project Group. The British journal of surgery. 1997 8;84:1130-5.

[42] van Gijn W, Marijnen CAM, Nagtegaal ID, Kranenbarg EMK, Putter H, Wiggers T, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. wwwthelancetcom/oncology. 2011;12:575-82. Available from: www.thelancet.com/oncology.

[43] Brændengen M, Tveit KM, Åke Berglund, Birkemeyer E, Frykholm G, Påhlman L, et al. Randomized phase III study comparing preoperative radiotherapy with chemoradiotherapy in nonresectable rectal cancer. Journal of Clinical Oncology. 2008;26:3687-94.

[44] Erlandsson J, Holm T, Pettersson D, Åke Berglund, Cedermark B, Radu C, et al. Optimal fractionation of preoperative radiotherapy and timing to surgery for rectal cancer (Stockholm III): a multicentre, randomised, non-blinded, phase 3, non-inferiority trial. The Lancet Oncology. 2017 3;18:336-46.

[45] Valentini V, Glimelius B, Haustermans K, Marijnen CAM, Rödel C, Gambacorta MA, et al. EURECCA consensus conference highlights about rectal cancer clinical management: The radiation oncologist's expert review. Radiotherapy and Oncology. 2014;110:195-8.

[46] Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rödel C, Cervantes A, et al.. Rectal cancer ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up on behalf of the ESMO Guidelines Committee; 2018. Available from: https://www.esmo.org/content/download/178876/3265837/1/Clinical-Practice-Guidelines-Slideset-Rectal-Cancer.pdf.

[47] Dijkstra EA, Hospers GAP, Kranenbarg EMK, Fleer J, Roodvoets AGH, Bahadoer RR, et al. Quality of life and late toxicity after short-course radiotherapy followed by chemotherapy or chemoradiotherapy for locally advanced rectal cancer – The RAPIDO trial. Radiotherapy and Oncology. 2022 6;171:69-76.

[48] Bahadoer RR, Dijkstra EA, van Etten B, Marijnen CAM, Putter H, Kranenbarg EMK, et al. Interpreting the RAPIDO trial: factors to consider – Authors' reply. The Lancet Oncology. 2021;22(3):e90-1. Available from: https://www.sciencedirect.com/science/article/pii/S1470204521000875.

[49] Byun HK, Koom WS. A practical review of watch-and-wait approach in rectal cancer. Department of Radiation Oncology; 2023.

[50] National Cancer Institute. Radiation Therapy to Treat Cancer [Online];. [updated 2019 Jan 08; cited 2023 Aug 29]. Available from: https://www.cancer.gov/about-cancer/treatment/types/radiation-therapy.

[51] National Cancer Institute. External Beam Radiation Therapy for Cancer [Online];. [updated 2018 May 01; cited 2023 Aug 30]. Available from: https://www.cancer.gov/about-cancer/treatment/types/radiation-therapy/external-beam.

[52] Khan FM, Gibbons JP. Khan's The Physics of Radiation Therapy. 5th ed. Philadelphia: Lippincott Williams Wilkins; 2014.

[53] Podgorsak EB. Radiation Oncology Physics: A Handbook for Teachers and Students. Vienna: International Atomic Energy Agency; 2005.

[54] Lilley JS. In: Nuclear Physics Principles and Applications. West Sussex: John Wiley Sons Ltd; 2001. .

[55] Evans RD. The Atomic Nucleus. New York: McGraw-Hill Company; 1955.

[56] Mott JHL, West NS. Essentials of Depth Dose Calculations for Clinical Oncologists. Elsevier Ltd; 2021.

[57] Sand O, Øystein V Sjaastad, Haug E. Menneskets fysiologi. 2nd ed. Oslo: Gyldendal; 2018.

[58] Stanford Medicine Children's Health. Mitochondrial Inheritance: Leber's Optic Atrophy [Online]. Stanford Medicine Children's Health;. [cited 2023 Dec 13]. Available from: https://www.stanfordchildrens.org/en/topic/default?id=mitochondrial-inheritance-lebers-optic-atrophy-90-P02131.

[59] Cleveland Clinic. DNA, Genes Chromosomes [Online]. Cleveland Clinic;. [updated 2022 May 20; cited 2024 Apr 08]. Available from: https://my.clevelandclinic.org/health/body/23064-dna-genes--chromosomes.

[60] Chang DS, Lasley FD, Das IJ, Mendonca MS, Dynlacht JR. Basic Radiotherapy Physics and Biology. 2nd ed. Springer International Publishing; 2021.

[61] van Leeuwen CM, Oei AL, Crezee J, Bel A, Franken NAP, Stalpers LJA, et al.. The alfa and beta of tumours: A review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies. BioMed Central Ltd.; 2018.

[62] Steel GG, Mcmillan TJ, Peacock JH. The 5rs of radiobiology. International Journal of Radiation Biology. 1989;56:1045-8.

[63] Withers HR. Four R's of radiotherapy. Advances in Radiation Biology. 1975;5 p. 241-247. Available from: https://inis.iaea.org/search/search.aspx?orig_q=RN:7228292.

[64] Kalender W. Computed tomography : fundamentals, system technology, image quality, applications. Erlangen: Publicis; 2011.

[65] Burnet NG, Thomas SJ, Burton KE, Jefferies SJ. Defining the tumour and target volumes for radiotherapy. Cancer Imaging. 2004;4:153-61.

[66] 3. Special Considerations regarding Absorbed-Dose and Dose–Volume Prescribing and Reporting in IMRT. Journal of the ICRU. 2010;10(1):27–40. Available from: https://doi.org/10.1093/jicru_ndq008.

[67] Herk MV. Errors and Margins in Radiotherapy. Seminars in Radiation Oncology. 2004;14:52-64.

[68] Stroom JC, Heijmen BJM. Geometrical uncertainties, radiotherapy planning margins, and the ICRU-62 report; 2002. Available from: www.elsevier.com/locate/radonline.

[69] Teague M. KV Imaging Systems for Linear Accelerators [Online]. Radiology Oncology Systems; 2015 Dec 19. [cited 2024 Apr 11]. Available from: https://www.oncologysystems.com/blog/kv-imaging-systems-for-linear-accelerators.

[70] Schulze R, Heil U, Groß D, Bruellmann DD, Dranischnikow E, Schwanecke U, et al.. Artefacts in CBCT: A review; 2011.

[71] Puvanasunthararajah S, Fontanarosa D, Wille ML, Camps SM. The application of metal artifact reduction methods on computed tomography scans for radiotherapy applications: A literature review. Journal of Applied Clinical Medical Physics. 2021 6;22:198-223.

[72] Zhang Y, Zhang L, Zhu XR, Lee AK, Chambers M, Dong L. Reducing metal artifacts in cone-beam CT images by preprocessing projection data. International Journal of Radiation Oncology Biology Physics. 2007 3;67:924-32.

[73] de Crevoisier R, Bayar MA, Pommier P, Muracciole X, Pêne F, Dudouet P, et al. Daily Versus Weekly Prostate Cancer Image Guided Radiation Therapy: Phase 3 Multicenter Randomized Trial. International Journal of Radiation Oncology Biology Physics. 2018 12;102:1420-9.

[74] Yan D, Vicini F, Wong J, Martinez A. Adaptive radiation therapy. Physics in Medicine Biology. 1997 jan;42(1):123. Available from: https://dx.doi.org/10.1088/0031-9155/42/1/008.

[75] Jong RD, Visser J, Crama KF, Wieringen NV, Wiersma J, Geijsen ED, et al. Dosimetric benefit of an adaptive treatment by means of plan selection for rectal cancer patients in both short and long course radiation therapy. Radiation Oncology. 2020 1;15.

[76] de Jong R, Visser J, van Wieringen N, Wiersma J, Geijsen D, Bel A. Feasibility of Conebeam CT-based online adaptive radiotherapy for neoadjuvant treatment of rectal cancer. Radiation Oncology. 2021 12;16.

[77] Gay HA, Barthold HJ, O'Meara E, Bosch WR, Naqa IE, Al-Lozi R, et al. Pelvic normal tissue contouring guidelines for radiation therapy: A radiation therapy oncology group consensus panel atlas. International Journal of Radiation Oncology Biology Physics. 2012 7;83.

[78] Myerson RJ, Garofalo MC, Naqa IE, Abrams RA, Apte A, Bosch WR, et al. Elective Clinical Target Volumes for Conformal Therapy in Anorectal Cancer:

A Radiation Therapy Oncology Group Consensus Panel Contouring Atlas. International Journal of Radiation Oncology Biology Physics. 2009 7;74:824-30.

[79] Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. BMC Medical Imaging. 2015 8;15.

[80] Sherer MV, Lin D, Elguindi S, Duke S, Tan LT, Cacicedo J, et al.. Metrics to evaluate the performance of auto-segmentation for radiation treatment planning: A critical review. Elsevier Ireland Ltd; 2021.

[81] Dice LR. Measures of the Amount of Ecologic Association Between Species; 1945.

[82] Boehringer AS, Sanaat A, Arabi H, Zaidi H. An active learning approach to train a deep learning algorithm for tumor segmentation from brain MR images. Insights into Imaging. 2023 12;14.

[83] Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993;15(9):850-63.

[84] Sharp G, Fritscher KD, Pekar V, Peroni M, Shusharina N, Veeraraghavan H, et al. Vision 20/20: Perspectives on automated image segmentation for radiotherapy. Medical Physics. 2014;41.

[85] Nuyttens JJ, Robertson JM, Yan DI, Martinez A. THE POSITION AND VOLUME OF THE SMALL BOWEL DURING ADJUVANT RADIATION THERAPY FOR RECTAL CANCER; 2001.

[86] Zhang F, Zhou M, Wang G, Li X, Yue L, Deng L, et al. Evaluation of bladder filling effects on the dose distribution during radiotherapy for cervical cancer based on daily CT images. Journal of Applied Clinical Medical Physics. 2023 11;24.

[87] National Institute of Diabetes and Digestive and Kidney Diseases. Your Digestive System How it Works [Online]. National Institutes of Health;. [updated 2017 Dec; cited 2024 Feb 29]. Available from: https://www.niddk.nih.gov/health-information/digestive-diseases/digestive-system-how-it-works.

[88] Holck P. tarmen [Online];. [updated 2018 Jul 29; cited 2023 Oct 10]. Store medisinske leksikon, snl.no. Available from: https://sml.snl.no/tarmen.

[89] Schuenke M, Schulte E, Schumacher U. THIEME Atlas of Anatomy. vol. 2. 2nd ed. Cass WA, Zeberg H, editors. Thieme; 2017.

[90] Sanguineti G, Little M, Endres EJ, Sormani MP, Parker BC. Comparison of three strategies to delineate the bowel for whole pelvis IMRT of prostate cancer. Radiotherapy and Oncology. 2008 7;88:95-101.

[91] Xu MJ, Kirk M, Zhai H, Lin LL. Bag and loop small bowel contouring strategies differentially estimate small bowel dose for post-hysterectomy women receiving pencil beam scanning proton therapy. Acta Oncologica. 2016 7;55:900-8.

[92] Banerjee R, Chakraborty S, Nygren I, Sinha R. Small Bowel Dose Parameters Predicting Grade 3 Acute Toxicity in Rectal Cancer Patients Treated With Neoadjuvant Chemoradiation: An Independent Validation Study Comparing Peritoneal Space Versus Small Bowel Loop Contouring Techniques. International Journal of Radiation Oncology*Biology*Physics. 2013;85(5):1225-31. Available from: https://www.sciencedirect.com/science/article/pii/S0360301612036462.

[93] Mövik L, Bäck A, Pettersson N. Impact of delineation errors on the estimated organ at risk dose and of dose errors on the normal tissue complication probability model. Medical Physics. 2023 3;50:1879-92.

[94] National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 [Online]. National Cancer Institute; 2010 Jun 14. [cited 2023 Dec 09]. Available from: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm.

[95] Baglan KL, Frazier RC, Yan DI, Huang RR, Martinez AA, Robertson JM. THE DOSE-VOLUME RELATIONSHIP OF ACUTE SMALL BOWEL TOXICITY FROM CONCURRENT 5-FU-BASED CHEMOTHERAPY AND RADIATION THERAPY FOR RECTAL CANCER; 2002.

[96] Robertson JM, Lockman D, Yan D, Wallace M. The Dose-Volume Relationship of Small Bowel Irradiation and Acute Grade 3 Diarrhea During Chemoradiotherapy for Rectal Cancer. International Journal of Radiation Oncology Biology Physics. 2008 2;70:413-8.

[97] Roeske JC, Bonta D, Mell LK, Lujan AE, Mundt AJ. A dosimetric analysis of acute gastrointestinal toxicity in women receiving intensity-modulated whole-pelvic radiation therapy. Radiotherapy and Oncology. 2003;69:201-7.

[98] Marks LB, Yorke ED, Jackson A, Haken RKT, Constine LS, Eisbruch A, et al. Use of Normal Tissue Complication Probability Models in the Clinic. International Journal of Radiation Oncology Biology Physics. 2010 3;76.

[99] Wasserthal J, Breit HC, Meyer MT, Pradella M, Hinck D, Sauter AW, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images • Content codes. Radiology: Artificial Intelligence. 2023;5. Available from: https://www.github.com/wasserth/TotalSegmentator.

[100] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods. 2021 2;18:203-11.

[101] Wasserthal J, Breit HC, Meyer MT, Pradella M, Hinck D, Sauter AW, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. Radiology: Artificial Intelligence. 2023;5(5):e230024. Available from: https://doi.org/10.1148/ryai.230024.

[102] Pálsson S, Cerri S, Poulsen HS, Urup T, Law I, Leemput KV. Predicting survival of glioblastoma from automatic whole-brain and tumor segmentation of MR images. Scientific Reports. 2022 12;12.

[103] Devore JL, Berk KN, Carlton MA. Modern Mathematical Statistics with Applications. 3rd ed. Springer Texts in Statistics. Cham: Springer International Publishing AG; 2021.

[104] Rohatgi A. WebPlotDigitizer; 2024. Available from: https://automeris.io/WebPlotDigitizer.html.

[105] Heijkoop ST, Langerak TR, Quint S, Mens JWM, Zolnay AG, Heijmen BJM, et al. Quantification of intra-fraction changes during radiotherapy of cervical cancer assessed with pre- and post-fraction Cone Beam CT scans. Radiotherapy and Oncology. 2015 12;117:536-41.

[106] Ahmad R, Hoogeman MS, Quint S, Mens JW, de Pree I, Heijmen BJM. Inter-fraction bladder filling variations and time trends for cervical cancer patients assessed with a portable 3-dimensional ultrasound bladder scanner. Radiotherapy and Oncology. 2008 11;89:172-9.

[107] Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Annals of Cardiac Anaesthesia. 2019 1;22:67-72.

[108] Perna L, Sini C, Cozzarini C, Agnello G, Cattaneo GM, Hysing LB, et al. Deformable registration-based segmentation of the bowel on Megavoltage CT during pelvic radiotherapy. Physica Medica. 2016 7;32:898-904.

[109] Alam S, Veeraraghavan H, Tringale K, Amoateng E, Subashi E, Wu AJ, et al. Inter- and intrafraction motion assessment and accumulated dose quantification of upper gastrointestinal organs during magnetic resonance-guided ablative radiation therapy of pancreas patients. Physics and Imaging in Radiation Oncology. 2022 1;21:54-61.

[110] Barten DLJ, van Kesteren Z, Laan JJ, Dassen MG, Westerveld GH, Pieters BR, et al.. Precision assessment of bowel motion quantification using 3D cine-MRI for radiotherapy. Institute of Physics; 2024.

[111] Gaudreault M, Siva S, Kron T, Hardcastle N. Assessing organ at risk position variation and its impact on delivered dose in kidney SABR. Radiation Oncology. 2022 12;17.

[112] Hysing LB, Kvinnsland Y, Lord H, Muren LP. Planning organ at risk volume margins for organ motion of the intestine. Radiotherapy and Oncology. 2006 9;80:349-54.

[113] Kvinnsland Y, Muren LP. The impact of organ motion on intestine doses and complication probabilities in radiotherapy of bladder cancer. Radiotherapy and Oncology. 2005 7;76:43-7.

[114] Nuyttens J, Robertson J, Yan D, Martinez A. The influence of small bowel motion on both a conventional three-field and intensity modulated radiation therapy (IMRT) for rectal cancer. Cancer/Radiothérapie. 2004 10;8:297-304.

[115] Smith GA, Dunlop A, Barnes H, Herbert T, Lawes R, Mohajer J, et al. Bladder filling in patients undergoing prostate radiotherapy on a MR-linac: The dosimetric impact. Technical Innovations and Patient Support in Radiation Oncology. 2022 3;21:41-5.

[116] Altman D, Zetterström J, López A, Pollack J, Nordenstam J, Mellgren A, et al. Effect of Hysterectomy on Bowel Function. Diseases of the Colon and Rectum. 2004 4;47:502-9.

[117] Gunnlaugsson A, Kjellén E, Nilsson P, Bendahl PO, Willner J, Johnsson A. Dose-volume relationships between enteritis and irradiated bowel volumes during 5-fluorouracil and oxaliplatin based chemoradiotherapy in locally advanced rectal cancer. Acta Oncologica. 2007;46:937-44.

[118] Varian Medical Systems. HyperSight™ - A Revolution in Resolution. Varian Medical Systems;. [cited 2024 Apr 27]. Available from: https://varian.widen.net/s/sd5kdw7wxv/hypersight_brochure_rad11063_oct2022.

[119] Varian Medical Systems. Varian Receives FDA 510(k) Clearance for TrueBeam and Edge Radiotherapy Systems Featuring the HyperSight Imaging Solution [Online]. Varian Medical Systems ; 2024 Feb 29. [cited 2024 Apr 27]. Available from: https://www.siemens-healthineers.com/press/releases/hypersight.

[120] Johansson A, Balter JM, Cao Y. Gastrointestinal 4D MRI with respiratory motion correction. Medical Physics. 2021 5;48:2521-7.

Appendix A

Justification for choices of statistical tests

Figure A6.1, A6.2 and A6.3 illustrate examples of density plots of intrafractional bowel motion and segmentation performances, normal Q-Q plots and individual value plots for volume differences between the manually delineated pre- and post-treatment bowels receiving dose for low and high dose levels, respectively, investigated prior to choosing statistical test.

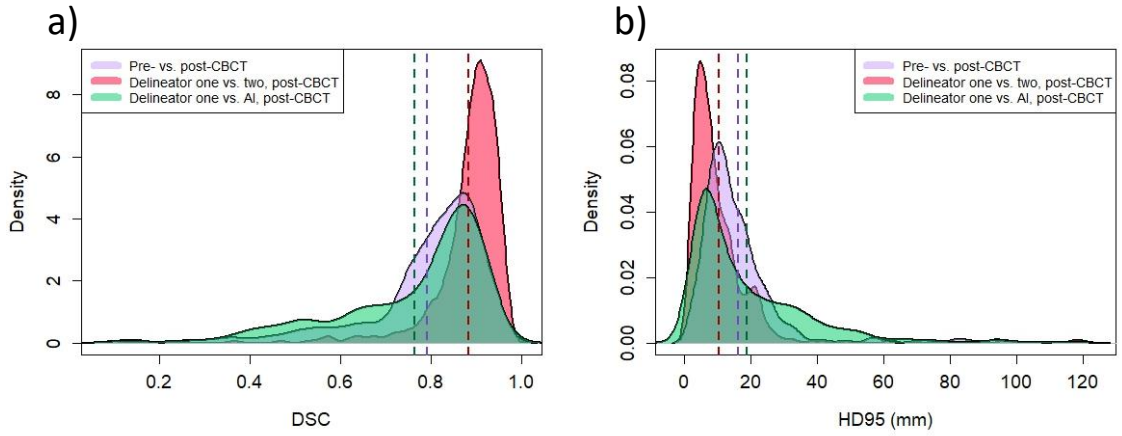

Figure 6.1: Density plots of a) Dice Similarity Coefficients (DSCs) and b) 95th Hausdorff Distances (HD95s) calculated between the manual bowel segmentations on pre- and post-CBCTs (purple), delineator one's and two's bowel segmentations on post-CBCTs (red) and delineator one's and TotalSegmentator's (AI) bowel segmentations on post-CBCTs (green). The values from all patients for fraction one are combined in the density plots. The mean values per segmentation comparison are plotted in dashed lines.

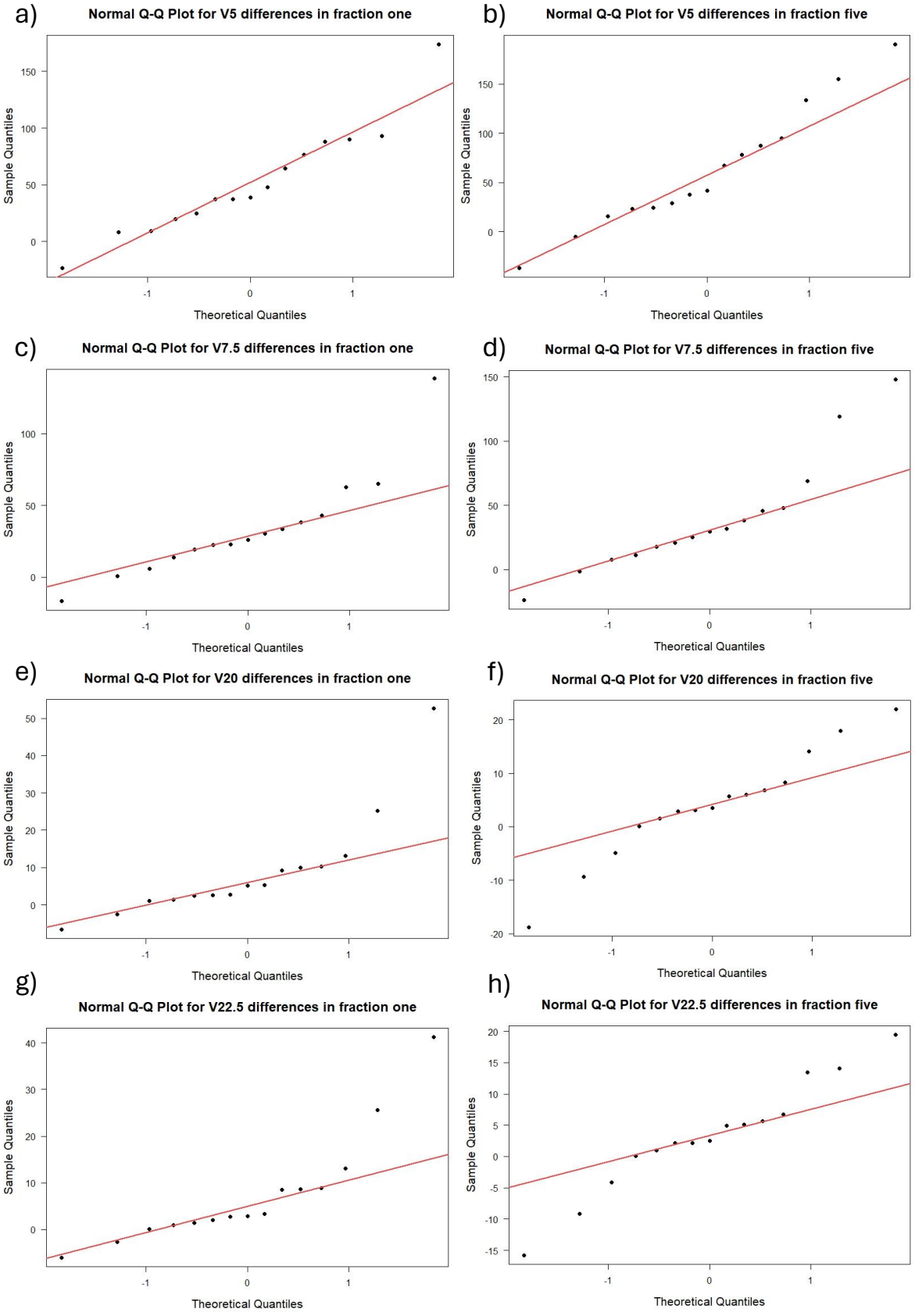

Figure 6.2: Normal Q-Q Plots of the volume differences between the manually delineated pre- and post-treatment bowels receiving dose ≥ a)-b) 5.0 Gy (V5), c)-d) 7.5 Gy (V7.5), e)-f) 20.0 Gy (V20) and g)-h) 22.5 Gy (V22.5) in fraction one and fraction five.

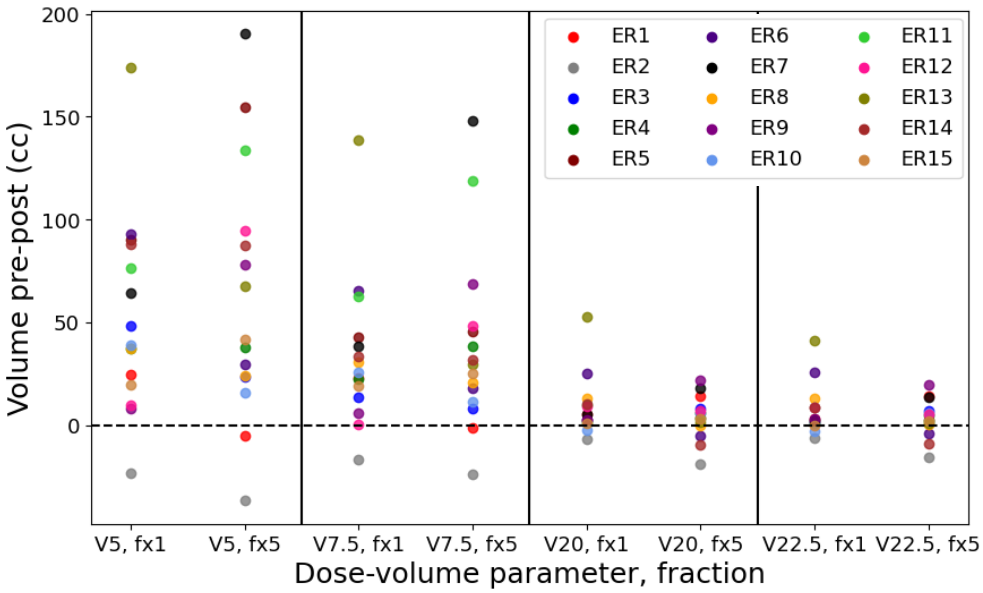

Figure 6.3: Individual value plot illustrating a point per patient representing the volume difference between the manually segmented pre- and post-treatment bowels receiving at least 5.0 Gy (V5), 7.5 Gy (V7.5), 20.0 Gy (V20) and 22.5 Gy (V22.5) in fraction one (fx1) and fraction five (fx5) separately.

Appendix B

Total segmented bowel volumes

Table B6.1 gives delineator one's, delineator two's, TS's (not post-processed) and TS's (post-processed) total segmented post-treatment bowel volume per patient.

Table 6.1: Overview of the total segmented post-treatment bowel volume (cc) of fraction one per patient, performed by the manual segmentators delineator one and two, and the Artificial Intelligence (AI) algorithm for automatic segmentations, TotalSegmentator. The post-processed AI segmentations refers to TS's combined small bowel and colon segmentations where single structures smaller than 0.50 cm^2 are removed and with an added outer 3D margin of 2 mm.

Total segmented bowel volume (cc) per patient, per segmentator				
Patient	**Delineator one**	**Delineator two**	**AI, post-processed**	**AI**
EthosRecti1	240.0	256.6	248.5	185.5
EthosRecti2	1477.4	1494.2	825.6	681.7
EthosRecti3	560.8	552.2	215.8	146.5
EthosRecti4	114.4	171.3	81.7	47.7
EthosRecti5	618.3	657.9	271.7	165.3
EthosRecti6	360.6	392.7	356.7	251.4
EthosRecti7	641.0	631.7	375.2	247.5
EthosRecti8	469.2	511.5	501.6	350.6
EthosRecti9	498.1	544.7	426.9	281.7
EthosRecti10	641.5	599.3	458.1	303.5
EthosRecti11	344.5	361.3	375.1	265.7
EthosRecti12	454.2	442.6	474.9	304.1
EthosRecti13	1239.5	1126.4	686.2	468.7
EthosRecti14	491.7	499.3	459.7	301.0
EthosRecti15	471.4	511.1	407.0	299.1
Average	574.8	583.5	411.0	286.7
Min	114.4	171.3	81.7	47.7
Max	1477.4	1494.2	852.6	681.7

Appendix C

Bladder volumes and relative intrafractional changes in bladder volumes

Table C6.2 gives the pre- and post-treatment bladder volumes in both fx1 and fx5 and the relative intrafractional changes in bladder volumes within both fx1 and fx5.

Table 6.2: Overview of the patient's pre- and post-treatment bladder volumes (cc) in fraction one and fraction five. The percentage volume change between the post- and pre-treatment bladders relative (rel.) to the pre-treatment bladder volumes are given (%).

Patient	**Bladder volume fraction one (cc)**			**Bladder volume fraction five (cc)**		
	Pre	**Post**	**Rel. volume difference (%)**	**Pre**	**Post**	**Rel. volume difference (%)**
ER1	332.6	360.1	8	356.8	388.8	9
ER2	213.3	203.1	−5	118.6	97.1	−18
ER3	290.0	273.9	−6	174.1	170.1	−2
ER4	310.7	356.0	15	41.6	42.0	1
ER5	80.8	94.7	17	83.8	109.8	31
ER6	387.9	412.6	6	120.5	106.0	−12
ER7	264.7	262.9	−1	206.4	209.3	1
ER8	52.4	55.3	6	80.0	86.9	−12
ER9	529.4	531.0	0	240.6	294.5	22
ER10	157.1	156.9	0	86.1	85.3	−1
ER11	110.6	101.4	−8	112.1	103.6	−8
ER12	219.4	211.7	−4	113.8	106.9	−6
ER13	224.0	245.4	10	288.1	281.8	−2
ER14	295.5	344.4	17	241.8	265.3	10
ER15	270.4	292.5	8	125.5	138.2	10

Appendix D

Segmentation performance in different genders

Figure D6.4 illustrates DSCs and HD95 values between delineator one's and TS's segmentation performances in male and female patients for the bladder, small bowel, colon and bowel. Figure D6.5 illustrates examples of two patients where TotalSegmentator confused the cervix-uterus with the bladder and the colon.

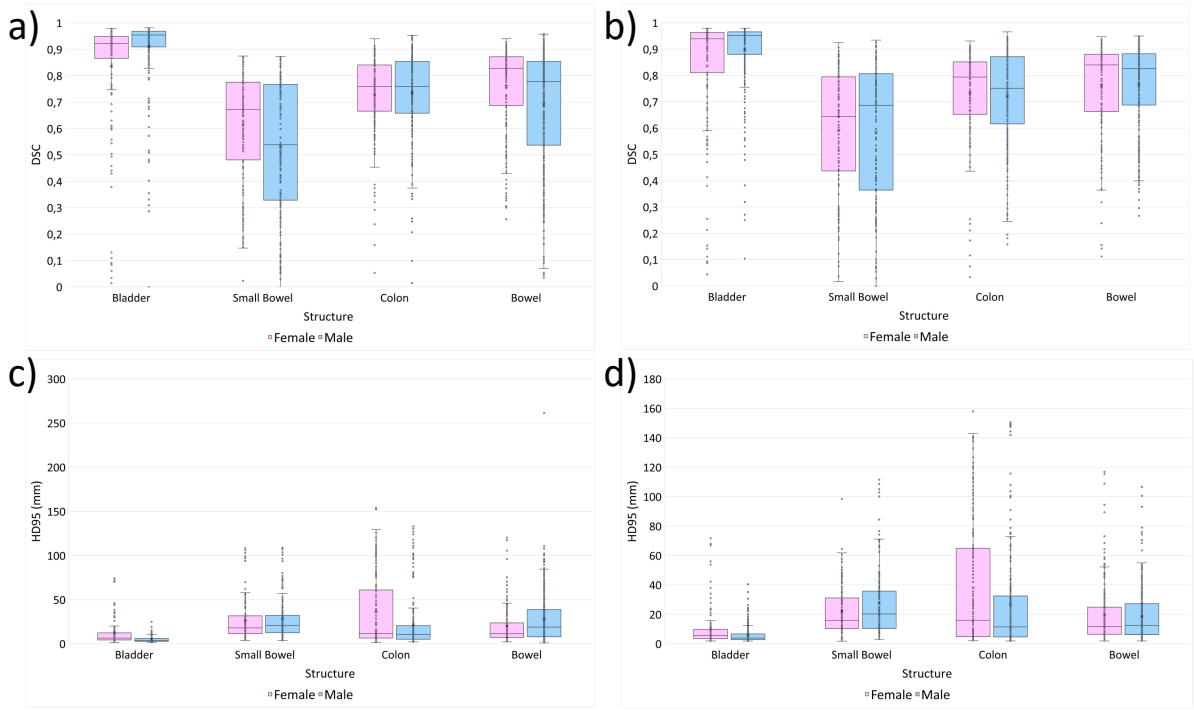

Figure 6.4: Segmentation performance plots quantifying TotalSegmentator's segmentation performance in different genders. Box plots of Dice Similarity Coefficients (DSCs) (a and b) and 95th percent Hausdorff Distances (HD95s) (c and d) for the bladder, small bowel, colon and bowel in male and female patients, calculated between the segmentations performed by delineator one and TS in pre-CBCTs (a and c) and post-CBCTs (b and d) of fraction one. Each box is delimited by the first and third quartile (Q1 and Q3) with a horizontal line indicating the median, a cross indicating the mean and whiskers extending outside the box. Outliers are illustrated as the data points outside the whiskers.

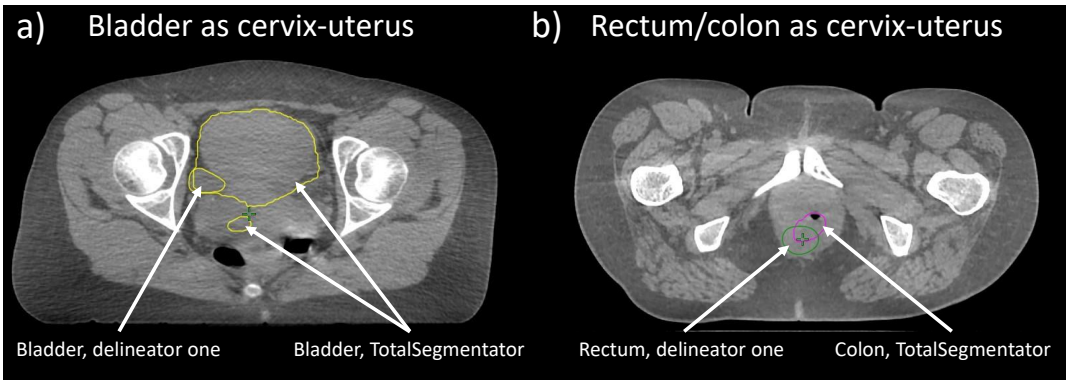

Figure 6.5: Comparisons of delineator one's and TotalSegmentator's segmentations in two female patients. a) TotalSegmentator included greater parts of the patient's cervix-uterus when segmenting the bladder. b) TotalSegmentator included some of the cervix-uterus when segmenting the colon. Here, delineator one's rectum corresponds with TotalSegmentator's colon because delineator one distinguished between colon and rectum while TotalSegmentator did not.

Appendix E

Adaptive dose plan

Figure E6.6 illustrates an example of an adaptive dose plan and simulated dose distribution within the depicted patient.

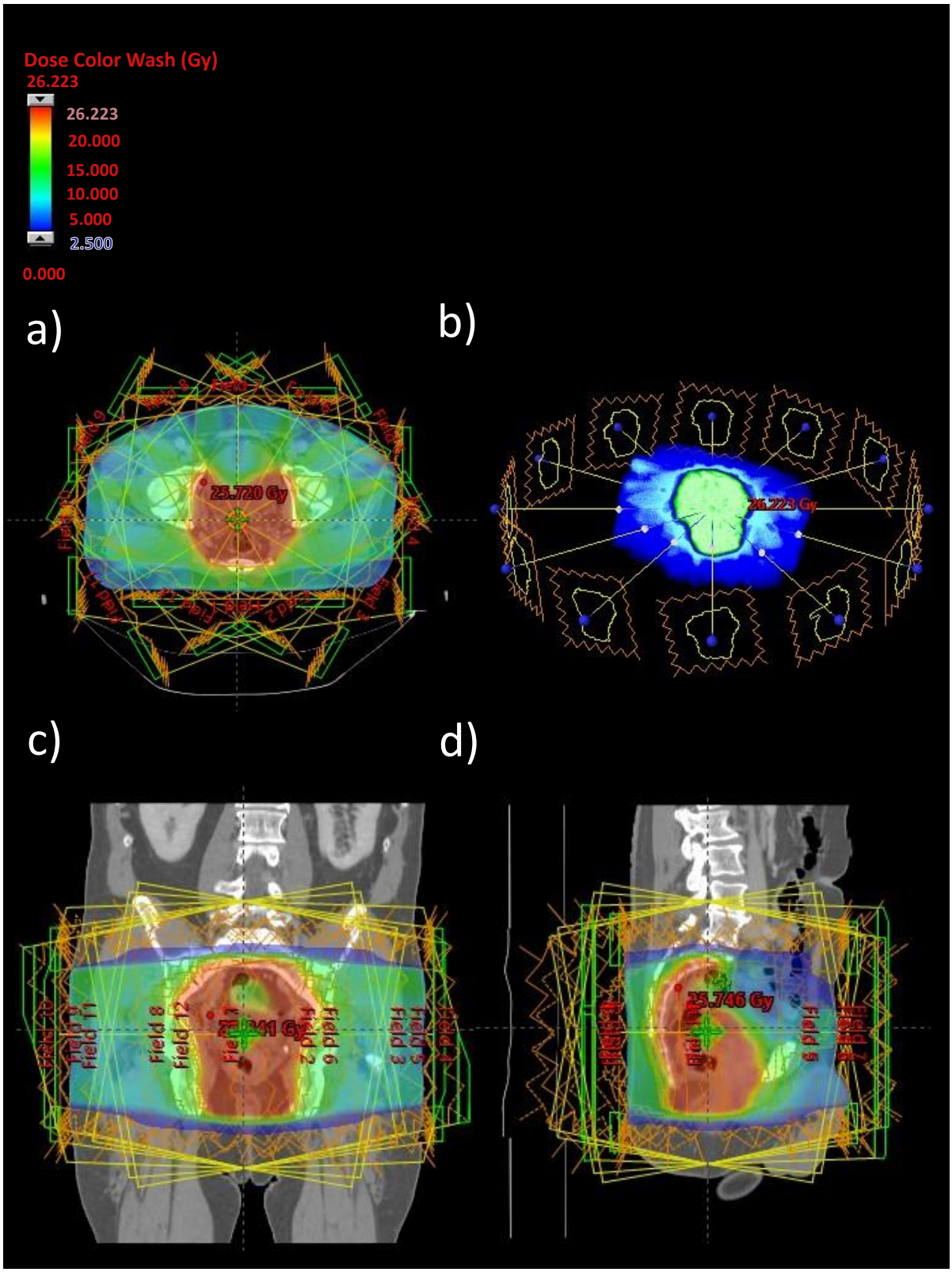

Figure 6.6: Dose plan for patient EthosRecti6, fraction five. a) Transversal slice, b) model view, c) frontal slice and d) sagittal slice illustrating 12 radiation fields and the simulated dose distribution in color wash explained by the color bar above.

Appendix F

MedFys abstract

The abstract submitted to MedFys2024 prior to oral presentation is shown below.

Background and Aim
In a recent clinical trial for rectal cancer, diarrhoea was the most frequent serious event during preoperative treatment [1]. Even though reduction of unnecessary dose to the bowel is part of the radiation treatment optimisation, bowel position can vary during treatment and the planned dose may not be representative for the dose delivered. This project aims to study bowel motion during fractions and investigate strategies to monitor and reduce the risk of toxicity.

Materials and Methods
In 15 rectal cancer patients, cone-beam CT scans (CBCTs) were acquired before and after treatment during five fractions each. A software for automatic segmentation, totalSegmentator[2][3], was run on all CBCT scans, and the bowel loop segmentations were then corrected to better fit the structures. A distinction was made between the small bowel and colon. In addition, the rectum was delineated to monitor target mobility. To quantify intra-fractional motion, Dice score and Hausdorff distance (95%) for small bowel, colon and rectum were calculated using an in-house Python script, comparing pre- and post-CBCT structures for all fractions for the first patient, and then for every first and last fraction. Further, a bowel motion model producing coverage probability maps will be established using a method previously developed by Hysing et al. [4].

Results and Conclusion
Preliminary results from the first two patients show total average Dice and Hausdorff score of 0.76 (sd ± 0.11) and 18.71 mm (sd ± 7.15 mm) for small bowel, 0.74 (sd ± 0.09) and 17.20 mm (sd ± 7.59 mm) for colon, and 0.89 (sd ± 0.04) and 4.23 mm (sd ± 1.24 mm) for rectum, respectively. The relatively moderate deviation from unity might be caused by delineation uncertainty, especially in areas of low CBCT quality due to air artefacts. Still, the qualitative image evaluation implies intra-fractional motion, particularly for bowel loops in close proximity to a changing bladder volume (Figure 1), which will be further investigated by the motion model.

Figure 1: Qualitative comparison of small bowel regions superior to the bladder. Coloured area represents non-overlapping area, whilst inside the dotted line represents the overlapping areas. The Dice and Hausdorff scores are given for each specific CBCT slice.

References
[1] Bahadoer RR et al., The Lancet Oncology. 2021;22(1):29-42.
[2] Wasserthal, J. et al., Radiology: Artificial Intelligence, 2023; 5:5.
[3] Isensee, F. et al., Nature methods, 2021;18(2), 203-211
[4] Hysing LB et al., Radiotherapy and Oncology, 2011; 100(3):407-411.