

Advanced Computed Tomography imaging in radiotherapy

Citation for published version (APA):

van der Heyden, B. (2020). *Advanced Computed Tomography imaging in radiotherapy*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20200615bh>

Document status and date:

Published: 01/01/2020

DOI:

[10.26481/dis.20200615bh](https://doi.org/10.26481/dis.20200615bh)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Advanced Computed Tomography imaging in radiotherapy

Brent van der Heyden

Advanced Computed Tomography imaging in radiotherapy

DISSERTATION

To obtain the degree of doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. Dr. Rianne M. Letschert,
in accordance with the decision of the Board of Deans,
to be defended in public
on Monday, 15th of June 2020 at 16:00 hours

by

Brent van der Heyden

SUPERVISOR

Prof. Dr. Frank Verhaegen

CO-SUPERVISORS

Dr. Mark A. J. Podesta

Dr. Gabriel P. Fonseca

ASSESSMENT COMMITTEE

Prof. Dr. Dirk de Ruyscher (chairman)

Prof. Dr. Dirk Verellen, VUB and Iridium kankernetwerk, Belgium

Dr. Brigitte Reniers, Hasselt University, Belgium

Dr. Ludwig J. Dubois

ISBN: 978-94-6380-723-4

Cover design: Ron Zijlmans

Lay-out: RON Graphic Power | www.ron.nu

Printing: ProefschriftMaken | www.proefschriftmaken.nl

© Brent van der Heyden, 2020

All rights are reserved. No part of this book may be reproduced, distributed, stored in a retrieval system, or transmitted in any form or by any means, without prior written permission of the author.

TABLE OF CONTENTS

Chapter 1	General introduction and outline	7
Chapter 2	VOXSI: A voxelized single- and dual-energy CT scenario generator for quantitative imaging	21
Chapter 3	Clinical evaluation of a novel CT image reconstruction algorithm for direct dose calculations: DirectDensity™	39
Chapter 4	A comparison study between single- and dual-energy CT density extraction methods for neurological proton Monte Carlo treatment planning	55
Chapter 5	Dual-energy CT for automatic organs-at-risk segmentation in brain-tumor patients using a multi-atlas and deep learning approach	79
Chapter 6	Modelling of the focal spot intensity distribution and the off-focal spot radiation in kilovoltage X-ray tubes for imaging	101
Chapter 7	A Monte Carlo based scatter removal method for non-isocentric c one-beam CT acquisitions using a deep convolutional autoencoder	121
Chapter 8	Automatic multi-atlas based organs-at-risk segmentation in mice	147
Chapter 9	Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network	163
Chapter 10	Summary, discussion and future perspectives	181
	Valorization addendum	211
	Acknowledgements	217
	Curriculum Vitae	221
	Publication list	225

The image features a vibrant watercolor background with a mix of warm colors including orange, yellow, and light blue. The colors are blended and layered, creating a soft, painterly effect. In the lower right quadrant, there is a large, white, hollow outline of the number '1'. The number is positioned vertically and has a clean, modern font style. The overall composition is abstract and artistic, with the number '1' standing out prominently against the colorful, textured background.

1

Chapter 1

General introduction and outline

Cancer, a growing health problem

Cancer is a general term for diseases in which endogenous cells divide in an uncontrolled way and may proliferate to other body parts through the blood and lymph systems. Considering the worldwide population growth [1, 2] and the increased life expectancy, the prevalence of cancer is a growing healthcare problem which has an escalating societal and economic impact worldwide. The Global Initiative for Cancer Registry Development reported 18.1 million diagnosed new cancer cases and 9.6 million cancer deaths worldwide in 2018 [3] and the numbers are rapidly growing, the International Agency for Research on Cancer (IARC) predicted that by 2030, ~26 million diagnosed cancer cases and 17 million cancer deaths will occur each year [4]. In 2017, the Netherlands Comprehensive Cancer Organization reported ~112 thousand new cancer cases and ~45 thousand cancer deaths nationally [5]. In 2018, the number of new cancer cases increased to 117 thousand [5].

Cancer treatment encompasses various invasive and non-invasive techniques such as surgery, chemotherapy, immunotherapy, radiotherapy, or more commonly a combination of these techniques to optimize treatment outcome. Surgery is an invasive operation where the cancer is excised from the body. Chemotherapy, a systemic therapy, uses drugs to kill or slow down the proliferation of cancer cells and immunotherapy, also a systemic therapy, is a biological treatment that aims to help the patient's immune system to fight the cancer itself. Radiotherapy can be invasive (e.g. intraoperative radiotherapy [6]) or non-invasive (external beam irradiation) and uses radiation to kill cancer cells. Various radiation types (Figure 1) are used to treat cancer depending on the clinical user case. Electrons (e^-) are considered for superficial treatments due to their limited penetration depth, ions and protons (p^+) are considered for irradiations that require a steep dose fall-off directly behind the treated volume, and megavoltage (MV) photons are the most commonly used particles to treat diverse tumor sites. Although most of the radiotherapy clinics treat patients with MV photons, in the Netherlands, several proton therapy centers became operational last years, e.g. HollandPTC in Delft, UMC PTC in Groningen, and Maastrro Protontherapie in Maastricht.

The role of imaging in radiotherapy

Medical imaging provides a visual representation of the human body's anatomy or function, and is essential for cancer screening, diagnosis, treatment planning, patient positioning, treatment verification and patient follow-up [7, 8]. During radiotherapy workflows many non-invasive or minimally invasive imaging modalities are clinically available, including Ultrasound (US) imaging, Magnetic Resonance (MR)

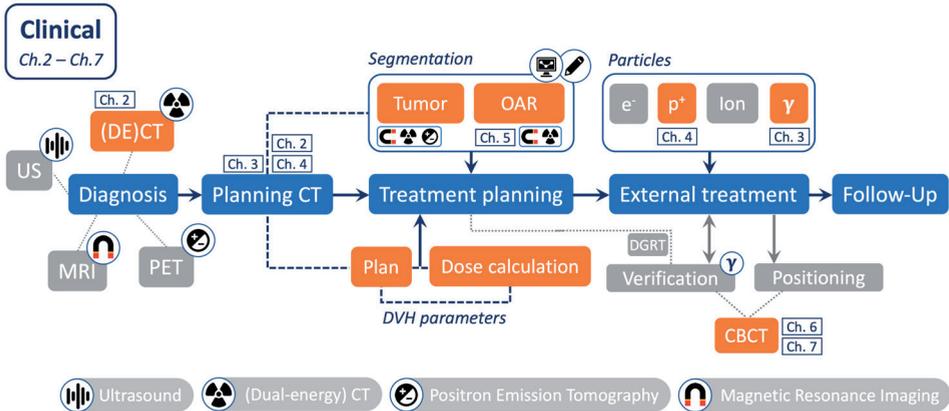


Figure 1. Schematic overview of the clinical radiotherapy workflow with indication of chapter numbers where this thesis made contributions. (OAR, organs-at-risk; DVH, dose volume histogram; DGRT, dose-guided radiotherapy; CBCT, Cone-Beam CT)

imaging, Computed Tomography (CT), Cone-Beam CT (CBCT), Positron Emission Tomography (PET), and portal beam imaging (Figure 1).

In MV photon treatments, modern treatment machines such as linear accelerators (linacs) have an on-board kV CBCT system and an on-board electronic portal imaging device (EPID) to visualize the patient anatomy and image the treatment beam to detect related errors [9, 10]. When calibrated for dose, the EPID can also act as a dosimeter to reconstruct the delivered treatment dose online and evaluate the actual (measured) dose distribution which can improve decision making, i.e. dose-guided radiotherapy (DGRT). All these imaging modalities play a vital role and are indispensable in a modern radiotherapy workflow. This thesis focusses primarily on tomography-based X-ray imaging, efforts to improve the CT and CBCT image quality that would affect the whole radiotherapy workflow, from imaging to treatment verification.

Computed Tomography (CT)

A single-energy (SE)CT scan is acquired by simultaneous rotation of an X-ray source and a detector around the patient to collect multiple X-ray projections. The projections are reconstructed into a volumetric image consisting of three-dimensional voxels. The voxel intensity ideally depends on the linear attenuation coefficient of its characterizing tissue, which is then converted to Hounsfield Units (HU) [11-13]. Patient CT images are used for radiotherapy treatment planning in which the CT-derived tissue properties (in HU) are converted to densities or stopping power ratios (SPR). The absolute CT number values (HU) are determined by several factors, such as the X-ray spectrum (kVp), the detector response, the tissue density, and the tissue composition. Changing one of these parameters, will lead to a CT scan of the same

geometry with different CT numbers. In dual-energy CT (DECT), which is a novel technology in radiotherapy [14-16], scanner settings are changed intentionally to obtain different CT numbers. These different CT numbers (HU) can be employed to derive additional tissue characteristics such as the effective atomic number (Z_{eff}) [17, 18], relative electron density (RED) [19, 20], and proton SPR relative to water [21-23], or to calculate pseudo-monoenergetic CT images (PMIs) [15, 24-28]. Extracting these parameters from DECT scans has been widely investigated to improve the accuracy of dose calculations in radiotherapy [29-33].

Cone-Beam Computed Tomography (CBCT)

The principle of CBCT imaging is very similar to the principle of fan-beam CT imaging. One main difference between both modalities is the divergent cone-shaped X-ray beam and the large flat-panel detector in CBCT which eliminates a moving table utilization for patient positioning during the acquisition. CBCT scanners have the advantage of being less expansive and being more compact, but as trade off, they provide a poorer image quality. The use of CBCT in precision radiotherapy is highly encouraged to verify the three-dimensional match between the planning position and the patient position prior to radiation [34]. Most CBCT systems used in photon radiotherapy are directly mounted on the linac. Because this is not feasible in proton therapy, alternatives such as in-room CT, robotic couch CBCT, or on-rail CBCT scanners are considered. Here, the medPhoton imaging ring system is one these novel robotic couch and on-rail CBCT scanners that provides non-isocentric imaging capabilities due to its X-ray source and detector that rotate independently from each other.

Furthermore, CBCT allows for the detection of significant anatomy changes between radiotherapy treatment fractions due to the patient treatment response or the disease progression. This is also relevant in preclinical research, where μ CBCT is used for the detection of anatomical changes in animals (e.g. **Chapter 9**). After the detection of clinically relevant anatomy changes (e.g. pleural effusion or atelectasis), adaptive radiotherapy treatment planning may be required to re-optimize the treatment plan [35, 36]. The inferior CBCT image quality compared to diagnostic CT modalities is a consequence of several effects; the flat-panel X-ray detector introduces image lag arising from charge trapping in the detector pixels [37], a much wider X-ray imaging beam leads to increased object scatter [38], and longer acquisition times lead to increased motion artifacts [39]. Object scatter majorly contributes to inferior image quality because it is disregarded in basic image reconstruction algorithms. Therefore, **Chapter 7** describes the development of a projection-based object scatter removal algorithm to improve CBCT image quality. Daily CBCT images are acquired in clinical radiotherapy routine, so methods to improve CBCT image quality would also allow for (improved) dose recalculation for photon and protons treatments [40-42].

Magnetic Resonance (MR)

A clinical MR scanner most commonly evaluates the nuclear spin properties of hydrogen atoms under the influence of a strong magnetic field to acquire anatomy images with superior soft-tissue contrast compared to CT. MR images cannot directly be used for dose calculations due to its arbitrary reconstruction units and its geometric image distortions than need thorough investigation [43, 44], hence, MR is in radiotherapy regularly combined with CT to evaluate tumor types and for the manual image segmentation of radiation sensitive tissues that need sparing during radiotherapy treatment. E.g. in **Chapter 4**, MR imaging was used along with CT to manually delineate the organs-at-risk (OARs) in brain cancer patients (e.g. hippocampi or chiasm).

Radiotherapy

During curative radiotherapy, high-energy ionizing radiation beams are used to induce cell death which inhibits both malignant tumor growth and the proliferation of metastasis. Ionizing treatment beams lead to breakage of chemical bonds and the release of highly reactive free radicals that cause DNA damage within cells. If the DNA damage is significant and rapid it will exceed the capabilities of the cellular repair mechanism and have the desired result of cell death. In radiotherapy treatments, the tumor is put in the crossfire of several radiation beams to concentrate the dose while simultaneously sparing the exposure of radiation sensitive cells in healthy tissues [45].

Innovative radiation technologies are an important part of the daily radiotherapy practice [46]. Generalized across the various tumor sites, the European Society for Radiotherapy and Oncology (ESTRO) estimated, as ‘ad hoc’ approach, that more than 50% of all cancer patients would benefit from innovative radiotherapy during the course of their disease [46, 47]. Consequently, annually ~59 thousand new cancer patients would benefit from improved radiotherapy technology in the Netherlands only.

Clinical radiotherapy workflow

A modern radiotherapy workflow consists of various stages following a cancer diagnosis (**Figure 1**). Before a radiotherapy treatment begins, a CT image is required to plan the treatment [48]. The quality of a radiotherapy treatment plan is evaluated using several metrics based on tumor dose and on the OAR dose, e.g. with dose-volume histograms (DVH). These doses and subsequently the DVH parameters can

only be derived and relied upon once the OAR and tumor boundaries are known. Commonly, medical physicists or radiotherapy technologists manually delineate the OARs and medical doctors manually delineate the tumor and possible metastases, all of which are subject to inter- and intra- observer variability [49, 50]. In many clinical scenarios it is challenging to distinguish tumor and OAR tissue from other neighboring soft tissues, so it is advantageous to include extra information from other imaging modalities such as MR [51] or PET [52].

In clinical treatment planning systems, the dose calculation algorithm requires material properties (e.g. relative electron density or mass density), this information is commonly obtained using a HU to material conversion curve which should be derived for each individual CT scanner and each available protocol. The HU to material curve accounts for many beam and object characteristics such as X-ray tube potential and patient size (**Chapter 3**). For brachytherapy dose calculations, many clinics make remarkable assumptions that the patient consists of water [53], even when more advanced DECT analyses showed that this assumption gives rise to drastically high dose distribution errors of maximum 20% [29].

Before the optimized plan delivery is started, commonly CBCT imaging is performed to verify the patient treatment position and to visually check anatomy changes that could cause the treatment dose to deviate from the calculated dose and negatively influence the treatment outcome. This kilovoltage CBCT can additionally be used for in vivo dosimetry, based on portal imaging in so called dose-guided radiotherapy (DGRT) [54].

Preclinical research workflow

Preclinical research assists the detailed investigation of several diseases in full detail. Reproducible small animal models could be used to intensively explore the induction and the growth of cancer, the disease progression, or the drug development [55-57]. Furthermore, small animal models are helpful to design optimal combined therapies (e.g. systemic therapy combined with radiotherapy). In radiotherapy studies, dedicated integrated preclinical imaging and irradiation cabinets can be used to mimic photon therapy treatments [58]. These preclinical imaging devices have much higher spatial μ CBCT voxel resolutions than clinical diagnostic CT scanners. μ CBCT voxel resolutions of $100 \times 100 \times 100 \mu\text{m}^3$ are not exceptional [59].

The preclinical radiotherapy workflow (**Figure 2**) deviates from the clinical workflow because the animal is under anesthesia while it is being irradiated. Imaging and irradiation are usually done in the same session, thus directly after the μ CBCT scan, the treatment must be planned while the animal is under anesthesia, including time consuming OAR segmentation, beam configuration planning, and

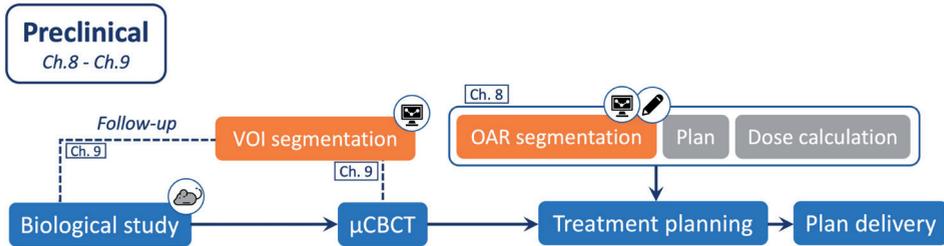


Figure 2. Schematic overview of a preclinical research workflow with indication of chapter numbers where this thesis made contributions (VOI, volume-of-interest; OAR, organ-at-risk; CBCT, micro Cone-Beam CT)

Monte Carlo dose calculations. All these workflow steps should not require more than ~30 minutes for the animal wellbeing [60], i.e. keeping the anesthesia duration to a minimum and avoid animal undercooling. In this thesis, **Chapter 8** describes the development of an automatic OAR image segmentation algorithm to accelerate the preclinical workflow prior to irradiation.

The submillimetric μ CBCT voxel sizes allow for the accurate volume-of-interest (VOI) determination (Figure 2). In longitudinal follow-up studies, imaging is often performed at specific timepoints to evaluate e.g. tumor growth [61]. To locate structure volumes, a time-consuming manual segmentation must be performed on multiple axial μ CBCT image planes. Considering the large animal cohorts and the large number of μ CBCT scans acquired at multiple timepoints, these studies result in a generous amount of data that must be analyzed manually. In **Chapter 9** of this thesis, a muscle mass determination algorithm is developed to analyze mice μ CBCT scans automatically and simultaneously provides the ability for longitudinal follow-up of cachexia in mice models.

Artificial Intelligence in radiotherapy

Computers in radiotherapy were most often used for dose calculations, image reconstructions, and data collection. The development of Artificial Intelligence (AI) techniques, such as machine learning and deep learning, have initiated a new era of big data analysis. Here, machine learning algorithms can potentially build mathematical models based on patterns in a numerical training dataset with minimal human assistance, e.g. in treatment outcome prediction. Deep learning for imaging in this context is based on artificial neural networks to classify, identify, or segment two- or three-dimensional image datasets; e.g. in OAR or tumor segmentation. In recent years, the use of these AI techniques has been investigated extensively in the literature to facilitate the clinical radiotherapy workflow. Techniques were examined to

help with, or even to take-over, the repetitive time-consuming human tasks. This thesis also investigated the use of AI in radiotherapy (**Chapter 5** and **Chapter 7**) and pioneered to apply AI in preclinical research (**Chapter 9**).

Thesis objectives

Imaging is used for screening, diagnosis, treatment planning, adaptive planning, patient treatment positioning, treatment verification, and therefore plays a fundamental role in the planning, validation and verification of radiotherapy treatments (**Figure 1**). This thesis investigates the impact level of novel tomography-based imaging modalities for radiotherapy, provides advanced imaging research software accessible to the community through software, and uses this gained knowledge to develop innovative methods to improve image quality and facilitate the (pre)clinical workflow.

Thesis outline

This thesis is divided in two parts. The first part focusses on imaging applications in clinical radiotherapy (**Chapters 2-7**) and the second part describes automatic image segmentation techniques in preclinical research (**Chapter 8** and **Chapter 9**). **Figure 1** and **Figure 2** depict where these thesis chapters are of interest in current clinical and preclinical workflows.

Clinical research

In this **Chapter 1**, common terminology used in this thesis is introduced, a general overview of the conventional radiotherapy workflow is provided, and the objectives of this thesis are covered.

Chapter 2 describes the development of a SECT and DECT simulator (VOXSI) for the simulation of realistic voxelized geometries with quantitative post-processing and analysis techniques. VOXSI is the first graphical user interface to perform DECT simulations and post-processing to calculate RED, Z_{eff} , PMIs, and material maps required for e.g. Monte Carlo radiotherapy dose calculations.

Chapter 3 discusses the unique market evaluation phase study and clinical commissioning of a novel, commercially available, image reconstruction algorithm for MV photon radiotherapy treatments that directly reconstructs CT images in relative electron densities instead of conventional Hounsfield Units.

Chapter 4 investigates differences of five different conversion methods for CT to relative electron density, in SECT or DECT for proton radiotherapy. First, the relative electron density accuracy was investigated in a phantom study for all methods. Then, the density images calculated by the different conversion methods were compared using four neurological cancer patients. Finally, dose and range differences for a proton treatment plan were investigated between the different conversion methods. This is the first study to compare multiple commercial and non-commercial CT number to relative electron density conversion methods for proton therapy dose calculations and aimed to provide guidelines, based on an extensive phantom study, on which algorithm to use in clinical routine wherein the patient ground truth is lacking

Chapter 5 analyses the quantitative and qualitative segmentation accuracy of an automatic multi-atlas and an automatic deep learning based OAR segmentation algorithm using DECT-derived PMIs of brain tumor patients. Both methods have been compared extensively in the literature for SECT images, but this study was the first to make DECT-based PMI reconstruction energy (keV) recommendations for both segmentation methods.

Chapter 6 models the novel medPhoton imaging ring CBCT modality in Monte Carlo and outlines a theoretical simulation study that models focal spot radiation and quantifies the effect of non-focal spot irradiation in a diagnostic X-ray tube caused by the backscattered electrons under the influence of the interelectrode electric field. This is today the most advanced diagnostic X-ray tube Monte Carlo simulation model compared to other published models because the cathode electron emission fluence is considered for two focal spot sizes together with the electric field.

Chapter 7 builds upon **Chapter 6**, and presents the development of a projection-based CBCT scatter removal algorithm using a neural network architecture that was trained on simulated Monte Carlo data of primary and scattered photon fluence distributions. The performance of the algorithm was first validated in CBCT simulations, and then in a retrospective study involving patient data. This is the first Monte Carlo driven CBCT scatter removal model validated on (non-)isocentric patient acquisitions and CBCT reconstructions.

Preclinical research

Chapter 8 investigates the first OAR image segmentation technique on CBCT relevant for the preclinical radiotherapy workflow. In this chapter, an atlas-based OAR segmentation method for mice was developed to accelerate pre-clinical treatment planning prior to irradiation.

Chapter 9 builds upon **Chapter 8**, and pioneered in the use of automatic CT image

Chapter 1 General introduction and outline

segmentation for other preclinical research fields, including cachexia research. In this chapter, a deep learning algorithm is developed to segment muscle groups of mice on micro CBCT images and to perform non-invasive longitudinal follow-up. Algorithm performance was tested and validated against muscle wet mass experiments.

Chapter 10 contains a summary and discussion of the work described in this thesis, its impact, and potential future avenues for study.

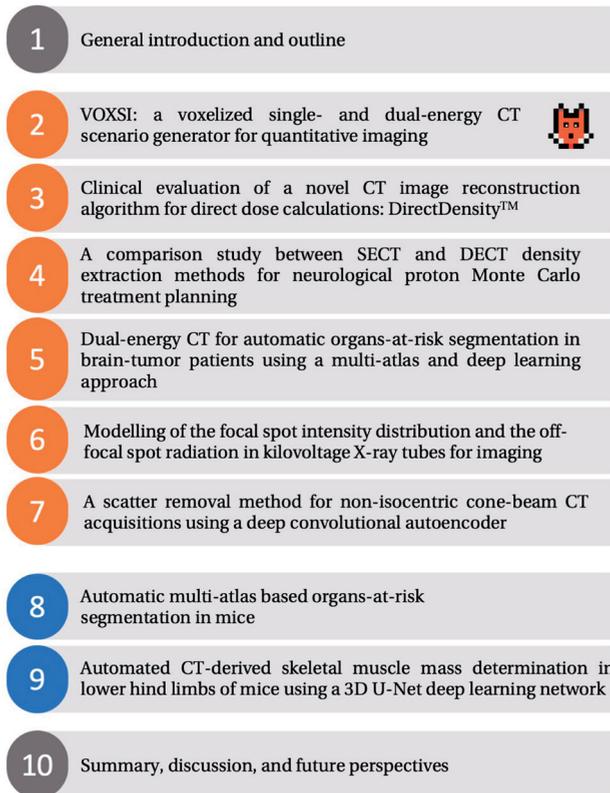


Figure 3. Schematic outline of this thesis. The orange field titles (Ch. 2-7) refer to clinical part 1 and the blue field titles (Ch. 8-9) refer to preclinical part 2 of this thesis.

References

1. Gerland, P., et al., *World population stabilization unlikely this century*. Science, 2014. **346**: p. 234-237.
2. Raftery, A.E., L. Alkema, and P. Gerland, *Bayesian Population Projections for the United Nations*. Stat Sci, 2014. **29**(1): p. 58-68.
3. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin, 2018. **68**(6): p. 394-424.
4. Ferlay, J., et al., *Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008*. Int J Cancer, 2010. **127**(12): p. 2893-917.
5. Nederland, I.K. *Cijfers over kanker*. 2019 [cited 2019; Available from: <https://www.cijfersoverkanker.nl>].
6. Paunesku, T. and G.E. Woloschak, *Future Directions of Intraoperative Radiation Therapy: A Brief Review*. Front Oncol, 2017. **7**: p. 300.
7. Fass, L., *Imaging and cancer: a review*. Mol Oncol, 2008. **2**(2): p. 115-52.
8. Alonzi, R. and P. Hoskin, *Functional imaging in clinical oncology: magnetic resonance imaging- and computerised tomography-based techniques*. Clin Oncol (R Coll Radiol), 2006. **18**(7): p. 555-70.
9. Langmack, K.A., *Portal imaging*. Br J Radiol, 2001. **74**(885): p. 789-804.
10. Srinivasan, K., M. Mohammadi, and J. Shepherd, *Applications of linac-mounted kilovoltage Cone-beam Computed Tomography in modern radiation therapy: A review*. Pol J Radiol, 2014. **79**: p. 181-93.
11. Schneider, U., E. Pedroni, and A. Lomax, *The calibration of CT Hounsfield units for radiotherapy treatment planning*. Phys Med Biol, 1996. **41**(1): p. 111-24.
12. Schneider, W., T. Bortfeld, and W. Schlegel, *Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions*. Phys Med Biol, 2000. **45**(2): p. 459-78.
13. Goma, C., I.P. Almeida, and F. Verhaegen, *Revisiting the single-energy CT calibration for proton therapy treatment planning: a critical look at the stoichiometric method*. Phys Med Biol, 2018. **63**(23): p. 235011.
14. van Elmpt, W., et al., *Dual energy CT in radiotherapy: Current applications and future outlook*. Radiother Oncol, 2016. **119**(1): p. 137-44.
15. Wohlfahrt, P., et al., *Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans*. Int J Radiat Oncol Biol Phys, 2017. **97**(2): p. 427-434.
16. Almeida, I.P., et al., *Dual-energy CT quantitative imaging: a comparison study between twin-beam and dual-source CT scanners*. Med Phys, 2017. **44**(1): p. 171-179.
17. Landry, G., et al., *Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients*. Phys Med Biol, 2013. **58**(19): p. 6851-66.
18. Saito, M. and S. Sagara, *A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body*. Med Phys, 2017. **44**(6): p. 2293-2303.
19. Saito, M., *Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship*. Med Phys, 2012. **39**(4): p. 2021-30.
20. Hunemohr, N., et al., *Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates*. Phys Med Biol, 2014. **59**(1): p. 83-96.
21. Paganetti, H., *Range uncertainties in proton therapy and the role of Monte Carlo simulations*. Phys Med Biol, 2012. **57**(11): p. R99-117.
22. ICRU., *Report 37: Stopping powers for electrons and protons*. .

23. Wohlfahrt, P., et al., *Evaluation of Stopping-Power Prediction by Dual- and Single-Energy Computed Tomography in an Anthropomorphic Ground-Truth Phantom*. Int J Radiat Oncol Biol Phys, 2018. **100**(1): p. 244-253.
24. Kuchenbecker, S., et al., *Dual energy CT: how well can pseudo-monochromatic imaging reduce metal artifacts?* Med Phys, 2015. **42**(2): p. 1023-36.
25. Yu, L., et al., *Virtual monochromatic imaging in dual-source dual-energy CT: radiation dose and image quality*. Med Phys, 2011. **38**(12): p. 6371-9.
26. Wichmann, J.L., et al., *Virtual monoenergetic dual-energy computed tomography: optimization of kiloelectron volt settings in head and neck cancer*. Invest Radiol, 2014. **49**(11): p. 735-41.
27. Michalak, G., et al., *Technical Note: Improved CT number stability across patient size using dual-energy CT virtual monoenergetic imaging*. Med Phys, 2016. **43**(1): p. 513.
28. Grant, K.L., et al., *Assessment of an advanced image-based technique to calculate virtual monoenergetic computed tomographic images from a dual-energy examination to improve contrast-to-noise ratio in examinations using iodinated contrast media*. Invest Radiol, 2014. **49**(9): p. 586-92.
29. Landry, G., et al., *Simulation study on potential accuracy gains from dual energy CT tissue segmentation for low-energy brachytherapy Monte Carlo dose calculations*. Phys Med Biol, 2011. **56**(19): p. 6257-78.
30. Cote, N., S. Bedwani, and J.F. Carrier, *Improved tissue assignment using dual-energy computed tomography in low-dose rate prostate brachytherapy for Monte Carlo dose calculation*. Med Phys, 2016. **43**(5): p. 2611.
31. Landry, G., et al., *Improved dose calculation accuracy for low energy brachytherapy by optimizing dual energy CT imaging protocols for noise reduction using sinogram affirmed iterative reconstruction*. Z Med Phys, 2016. **26**(1): p. 75-87.
32. Almeida, I.P., et al., *Monte Carlo proton dose calculations using a radiotherapy specific dual-energy CT scanner for tissue segmentation and range assessment*. Phys Med Biol, 2018. **63**(11): p. 115008.
33. Vaniqui, A., et al., *The impact of dual energy CT imaging on dose calculations for pre-clinical studies*. Radiat Oncol, 2017. **12**(1): p. 181.
34. Bell, K., et al., *Image guidance and positioning accuracy in clinical practice: influence of positioning errors and imaging dose on the real dose distribution for head and neck cancer treatment*. Radiat Oncol, 2018. **13**(1): p. 190.
35. Duffton, A., et al., *An assessment of cone beam CT in the adaptive radiotherapy planning process for non-small-cell lung cancer patients*. Br J Radiol, 2016. **89**(1062): p. 20150492.
36. van den Bosch, M., et al., *Automatic selection of lung cancer patients for adaptive radiotherapy using cone-beam CT imaging*. Physics and Imaging in Radiation Oncology, 2017. **1**: p. 21-27.
37. Siewerdsen, J.H. and D.A. Jaffray, *Cone-beam computed tomography with a flat-panel imager: effects of image lag*. Med Phys, 1999. **26**(12): p. 2635-47.
38. Siewerdsen, J.H. and D.A. Jaffray, *Cone-beam computed tomography with a flat-panel imager: magnitude and effects of X-ray scatter*. Med Phys, 2001. **28**(2): p. 220-31.
39. Nardi, C., et al., *Motion artefacts in cone beam CT: an in vitro study about the effects on the images*. Br J Radiol, 2016. **89**(1058): p. 20150687.
40. Giacometti, V., et al., *An evaluation of techniques for dose calculation on cone beam computed tomography*. Br J Radiol, 2019. **92**(1096): p. 20180383.
41. Park, Y.K., et al., *Proton dose calculation on scatter-corrected CBCT image: Feasibility study for adaptive proton therapy*. Med Phys, 2015. **42**(8): p. 4449-59.

42. Arai, K., et al., *Feasibility of CBCT-based proton dose calculation using a histogram-matching algorithm in proton beam therapy*. *Phys Med*, 2017. **33**: p. 68-76.
43. Adjeiwaah, M., et al., *Quantifying the Effect of 3T Magnetic Resonance Imaging Residual System Distortions and Patient-Induced Susceptibility Distortions on Radiation Therapy Treatment Planning for Prostate Cancer*. *Int J Radiat Oncol Biol Phys*, 2018. **100**(2): p. 317-324.
44. Adjeiwaah, M., et al., *Dosimetric Impact of MRI Distortions: A Study on Head and Neck Cancers*. *Int J Radiat Oncol Biol Phys*, 2019. **103**(4): p. 994-1003.
45. Baskar, R., et al., *Biological response of cancer cells to radiation treatment*. *Front Mol Biosci*, 2014. **1**: p. 24.
46. Lievens, Y. and C. Grau, *Health economics in radiation oncology: introducing the ESTRO HERO project*. *Radiother Oncol*, 2012. **103**(1): p. 109-12.
47. Borrás, J.M., et al., *The impact of cancer incidence and stage on optimal utilization of radiotherapy: Methodology of a population based analysis by the ESTRO-HERO project*. *Radiother Oncol*, 2015. **116**(1): p. 45-50.
48. Grau, C., et al., *Radiotherapy equipment and departments in the European countries: final results from the ESTRO-HERO survey*. *Radiother Oncol*, 2014. **112**(2): p. 155-64.
49. Segedin, B. and P. Petric, *Uncertainties in target volume delineation in radiotherapy - are they relevant and what can we do about them?* *Radiol Oncol*, 2016. **50**(3): p. 254-62.
50. Vinod, S.K., et al., *Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies*. *Radiother Oncol*, 2016. **121**(2): p. 169-179.
51. Eekers, D.B., et al., *The EPTN consensus-based atlas for CT- and MR-based contouring in neuro-oncology*. *Radiother Oncol*, 2018. **128**(1): p. 37-43.
52. Konert, T., et al., *PET/CT imaging for target volume delineation in curative intent radiotherapy of non-small cell lung cancer: IAEA consensus report 2014*. *Radiother Oncol*, 2015. **116**(1): p. 27-34.
53. Rivard, M.J., et al., *Update of AAPM Task Group No. 43 Report: A revised AAPM protocol for brachytherapy dose calculations*. *Med Phys*, 2004. **31**(3): p. 633-74.
54. van Rooijen, D.C., et al., *Dose-guided radiotherapy: potential benefit of online dose recalculation for stereotactic lung irradiation in patients with non-small-cell lung cancer*. *Int J Radiat Oncol Biol Phys*, 2012. **83**(4): p. e557-62.
55. Zuberi, A. and C. Lutz, *Mouse Models for Drug Discovery. Can New Tools and Technology Improve Translational Power?* *ILAR J*, 2016. **57**(2): p. 178-185.
56. Koontz, B.F., F. Verhaegen, and D. De Ruyscher, *Tumour and normal tissue radiobiology in mouse models: how close are mice to mini-humans?* *Br J Radiol*, 2017. **90**(1069): p. 20160441.
57. Butterworth, K.T., K.M. Prise, and F. Verhaegen, *Small animal image-guided radiotherapy: status, considerations and potential for translational impact*. *Br J Radiol*, 2015. **88**(1045): p. 20140634.
58. Verhaegen, F., P. Granton, and E. Tryggestad, *Small animal radiotherapy research platforms*. *Phys Med Biol*, 2011. **56**(12): p. R55-83.
59. Vaniqui, A., et al., *The effect of different image reconstruction techniques on pre-clinical quantitative imaging and dual-energy CT*. *Br J Radiol*, 2019. **92**(1095): p. 20180447.
60. Balvert, M., et al., *A framework for inverse planning of beam-on times for 3D small animal radiotherapy using interactive multi-objective optimisation*. *Phys Med Biol*, 2015. **60**(14): p. 5681-98.
61. Sosa Iglesias, V., et al., *An orthotopic non-small cell lung cancer model for image-guided small animal radiotherapy platforms*. *Br J Radiol*, 2019. **92**(1095): p. 20180476.

2

Chapter 2

VOXSI: A voxelized single- and dual-energy CT scenario generator for quantitative imaging

Brent van der Heyden, Lotte EJR Schyns, Mark Podesta, Ana Vaniqui,
Isabel P Almeida, Guillaume Landry, and Frank Verhaegen.

Published in

Physics and Imaging in Radiation Oncology (2018)

<https://doi.org/10.1016/j.phro.2018.05.004>

Abstract

Background and Purpose: Dedicated CT simulation models have the potential to investigate several acquisition, reconstruction, or post-processing parameters without giving any radiation dose to patients. In this chapter, a software program was developed for the simulation and the analysis of single-energy and dual-energy CT images. Simulation and analysis functionalities of the software are described in this thesis chapter.

Materials and Methods: In the software, named *VOXSI* (VOXelized CT Simulator), the X-ray source, user specified simulation geometry, CT setup and the detector energy response can be varied. CT image reconstructions can be performed with an implementation of the ASTRA toolbox. In the dual-energy CT post processing toolkit, GUI tools are provided to calculate effective atomic number, relative electron density, pseudo-monoenergetic, and material map images. Quantitative CT number validation, based on the RMI 467 tissue characterization phantom (Gammex), was performed between experimental and simulated CT scans at three different X-ray tube potentials (80, 120, and 140 kVp) with the SOMATOM Confidence RT Pro CT scanner (Siemens Healthineers).

Results: Overall, a good agreement was found for the mean CT numbers of the RMI 467 inserts. For all energies, the maximum difference in CT numbers between experimental and simulated data was below 17 HU for the soft tissues and below 48 HU for the osseous tissues.

Conclusion: The software's simulation algorithm showed a good agreement between the CT measurements and CT simulations of the RMI 467 phantom at different X-ray tube potentials. The capabilities of the software are demonstrated by an elaborated dual-energy CT research example.

Conflict of Interest

This work was partially performed with research funding from Varian Medical Systems (Palo Alto, USA).

Introduction

X-ray imaging is widely used to acquire images in many fields, such as oncology and radiotherapy, where imaging plays a crucial role in diagnosing and treating a wide range of cancerous diseases. An image modality such as single-energy (SE) computed tomography (CT) is frequently used for radiology and radiotherapy treatment planning [1]. Besides conventional SECT imaging, interest in dual-energy CT (DECT) for radiology and radiotherapy is increasing [2]. A growing body of literature suggests that DECT is beneficial for subjective and objective image quality, soft tissue characterization, projection or image based metal artifact reduction, brachytherapy dose calculations, particle therapy dose calculations, and small animal radiotherapy dose calculations [3-17].

Multiple aspects of CT imaging should be investigated to establish an optimal scanning protocol for each specific treatment site and patient size. Unfortunately, changing hardware components, or varying imaging parameters is not always possible in a clinical workflow. A dedicated CT simulation model has the potential to investigate scanning protocols, reconstruction algorithms, and beam hardening effects without giving radiation dose to patients or requiring the development of prototypes and acquisition of phantom scans.

In the literature, simulation tools for different imaging modalities were published. For example, tools were presented to simulate planar X-ray images [18-20]. Ay and Zaidi developed a fan-beam CT and cone-beam CT X-ray CT simulator of mathematical volumes based on Monte Carlo (MC) simulations [21]. Other researchers applied a ray-tracing approach algorithm to simulate fan-beam CT images [22, 23]. Segars *et al.* developed a parallel-beam, fan-beam, and cone-beam CT projector to simulate realistic humanoid anatomies from mesh volumes, without an analysis toolkit that provides tools to process the simulated data [24, 25]. Another simulation tool, *ImaSim* simulates planar kilo- and megavoltage, fan-beam CT, and cone-beam CT (without scatter) images of simulation geometries formed by combining only a few mathematical volumes [26].

Current CT simulation models use either MC simulations or a ray-tracing approach. MC simulates individual particle interactions and requires much more processing power compared to the ray-tracing alternative which ignores individual photon interactions to speed up the calculation time.

In this thesis chapter, a new fan-beam CT scenario generator accessible via a graphical user interface (GUI) is described: VOXSI. The simulation software has the flexibility to use a variety of voxelized geometries, and to modify X-ray spectra, geometrical setup, detector responses, and CT image reconstruction parameters. The analysis software toolkit includes DECT analysis capabilities to examine relevant parameters such as the relative electron density (RED), effective atomic number

(Z_{eff}), pseudo-monoenergetic energy images (PMIs), and material maps. In material maps, unique voxel-specific material flags are assigned to obtain a spatial material distribution. Compared with other available CT simulation packages, *VOXSI* allows the user to perform a multi-threaded CT simulation, from voxelized user defined geometries, while many simulation and reconstruction parameters can be varied. No other simulation packages are currently available that allow the user to perform different DECT analyses. For validation of our simulation model, the model was compared with measurements at three different X-ray tube potentials.

Materials and Methods

A schematic overview of the different tasks and toolkits embedded in the software are depicted in [Figure 1](#). The simulation software, named *VOXSI* (VOXelized CT Simulator), was developed in MATLAB® (The MathWorks Inc, Natick, MA) and runs on 64-bit Windows and Linux computers.

X-ray source

As X-ray spectrum generator, *SpekCalc* was implemented in *VOXSI*'s GUI [27]. *SpekCalc* is a validated simulation software package which calculates filtered polychromatic X-ray spectra (40-300 kVp) for X-ray tubes with a tungsten anode via a deterministic/MC model. The model calculates the bremsstrahlung spectrum in combination with a MC pre-calculated energy and depth distribution of electrons in a tungsten target [28, 29]. Additionally, measured spectra, spectra simulated in an external package, monoenergetic photons, and characteristic gamma rays emitted by radionuclides can be used as input of *VOXSI*'s CT simulation algorithm.

Voxelized geometry

VOXSI was designed such that only voxelized geometries are used as input for the CT simulation algorithm, e.g. DICOM CT scans or user specified voxelized geometries. Delineation tools, provided in the GUI, such as a brush (draw/erase) and intensity-based region growing help the user to assign materials in the voxels of a DICOM CT scan. The user specified geometries should have unique integer values assigned to every voxel. In *VOXSI*, these unique integer values can be linked to a material in the material database. In a recent update, radiotherapy structures that were segmented in e.g. a radiotherapy treatment planning system can be imported to create more advanced three-dimensional material maps.

The 3D humanoid XCAT phantom, which also has 4D capabilities, can serve as a realistic and flexible voxelized phantom [25]. The activity mode of the XCAT phantom produces a humanoid phantom with unique organ labels, defined in the XCAT

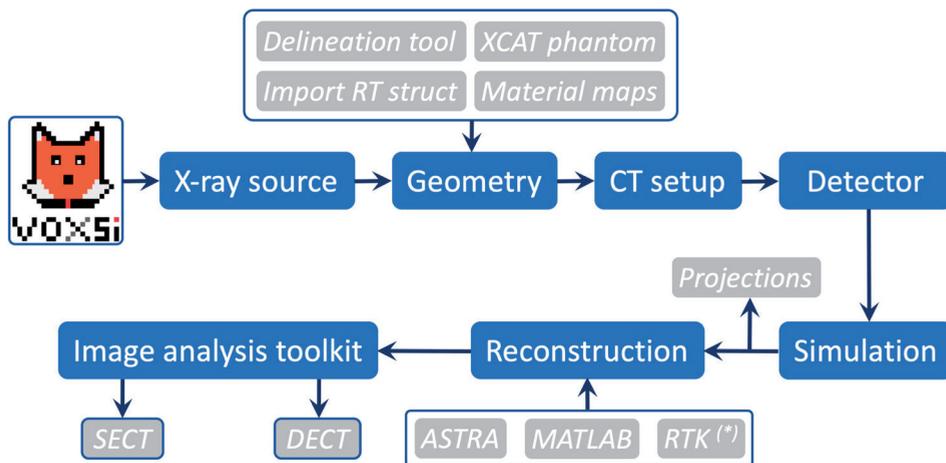


Figure 1. VOXSI flowchart. The X-ray source, simulation geometry, CT setup, and the detector energy response must be defined before the simulation start. After simulation, the CT projections are exportable, or can be reconstructed using the implemented ASTRA toolbox or MATLAB's FBP algorithm. Alternatively^(*), the open-source image reconstruction toolkit (RTK) can be used externally (not implemented in the GUI) to reconstruct the CT projections. The SECT and DECT image analysis toolkits provide tools to analyze the reconstructed CT images. Snapshots of VOXSI's interface are added in [SupplementaryMaterial A](#). (Abbreviations: XCAT, 4D extended cardiac-torso phantom; SECT, single-energy CT; DECT, dual-energy CT)

parameter file, which can be loaded in VOXSI's geometry editor. This methodology is also applicable on the 4D mouse and rat whole body phantoms: MOBY and ROBY [30].

CT setup and detector response

VOXSI simulates a non-helical fan-beam CT setup, although clinical CT scanners often scan in helical mode. A non-helical approach is assumed because a voxelized geometry consists of discrete slices. Different fan-beam setup parameters can be varied in the software; source to origin distance, detector to origin distance, fan-beam angle, curved or planar detector, number of detectors, and the number of CT projections.

An essential input of VOXSI is the detector energy response, defined as the ratio of the absorbed and incident photon energy. The detector response curve can often be obtained from the manufacturers by request. Alternatively, the detector response curve can be obtained by performing MC simulations in external software using the detector specifications [31, 32]. Whichever way it was obtained, the requested, pre-calculated, or even measured detector energy response curve can be loaded in VOXSI. As standard, an ideal detector response with a 100% absorption for all incident photon energies, is provided in the software.

Simulation and CT image reconstruction

A multithreaded ray-tracing approach was implemented in *VOXSI* to reduce calculation times. E.g. 45 seconds of computing time was needed to simulate a 512x512x1 geometry, a 120 kVp X-ray spectrum, 1024 detectors, and 780 projections, on two Intel® Xeon® 2.67 GHz X550 processors.

VOXSI utilizes the total mass attenuation coefficients, grouping all main three photon interaction mechanisms, photoelectric absorption, Rayleigh scatter, and Compton scatter, of the attenuating material that was assigned to the intersecting voxel. The polychromatic X-ray attenuation in each voxel is calculated by using the total mass attenuation coefficient as a function of the photon energy and the radiological path length, according to Lambert-Beer's attenuation law. The radiological path length through the voxel grid is calculated according to the algorithm described in [31]. The mass attenuation coefficients of a specific material can be calculated from the photon cross section database (XCOM version 3.1) [33]. Bowtie filters of different materials, described by their focal spot distance, thickness, and curvature, can be added to *VOXSI*'s CT simulation model.

To correct for cupping artifacts in polychromatic CT simulations, a basic beam hardening correction algorithm can be applied on the raw CT projection data. First, the algorithm assumes that the original simulation geometry consists of liquid water. Then, ideal projections (ignoring beam hardening) and spectrum hardened projections are calculated using the polychromatic X-ray spectrum, the detector energy response, and the radiological path length through the simulated water geometry. Finally, the ratio between the hardened and the ideal projection data was used to correct the projection data of the original simulation geometry for beam hardening artifacts. Because the raw projection data can be extracted from *VOXSI*, the user has the ability to apply their own beam hardening correction algorithm in external software.

An open source platform for image reconstruction, the ASTRA toolbox was implemented in *VOXSI* to reconstruct the simulated raw CT projection data [34]. The fan-beam projection data can be reconstructed with filtered back projection (FBP), a simultaneous iterative reconstruction technique (SIRT), a simultaneous algebraic reconstruction technique (SART), and a conjugate gradient least squares (CGLS) algorithm. These reconstruction techniques are based on a planar projector, while the projection data is simulated with either a curved or a planar projector. Therefore, the fan-beam projection data is converted to parallel-beam projection data by applying a rebinning procedure. Poisson noise can be added in the projection domain to simulate CT scans acquired at a lower imaging dose.

Single-energy CT and dual-energy CT image analysis toolkit

VOXSI has an implemented SECT and DECT analysis toolkit. The SECT analysis toolkit allows the user to calculate various image metrics (e.g. minimum, maximum, and mean Hounsfield Units ± 1 standard deviation) in regions of interest defined by the material segmentation. These metrics could afterward be exported in Excel format (Windows operating systems only). The more extended DECT analysis toolkit (GUI) allows the user to calculate RED images according to Saito's method, Z_{eff} images according to both Landry's method (2013) and Saito and Sagara's method, PMIs following the principles in Yu et al., and material maps according to method described in Landry et al (2011) [35-39]. Simulated low energy CT_L and high energy CT_H image data of the Gammex RMI 467 tissue characterization phantom (Gammex, Middleton, WI) inserts could be used to obtain the fit parameters.

Saito's method calculates RED images based on a linear fit procedure, using energy-subtracted CT numbers ΔHU , calculated by a weighted subtraction of CT_L and CT_H [40]. In [Equation 1](#) and [Equation 2](#), a, b, and α are fit parameters.

$$\Delta HU = (1 + \alpha) \cdot CT_H - \alpha \cdot CT_L \quad (\text{Eq. 1})$$

$$\text{RED} = a \cdot \frac{\Delta HU}{1000} + b \quad (\text{Eq. 2})$$

Landry's method calculates $Z_{\text{eff}}^{(1)}$ images based on a fit procedure. [Equation 3](#) shows the fitting model between the ratio of reduced CT numbers $v_{L/H} \equiv \frac{CT_{L, \text{RED}}}{1000} + 1$, for CT_L or CT_H , which involves several fit parameters; A, B, C, D, E, F, m, and n.

$$v_{\text{ratio}} = \frac{v_L}{v_H} = \frac{A + B \cdot Z_{\text{eff}}^{n-1} + C \cdot Z_{\text{eff}}^{m-1}}{D + E \cdot Z_{\text{eff}}^{n-1} + F \cdot Z_{\text{eff}}^{m-1}} \quad (\text{Eq. 3})$$

Quantitative CT validation

To validate VOXSI's simulation algorithm quantitatively, the RMI 467 phantom with its certified tissue mimicking inserts was scanned at three X-ray tube potentials (80, 120, and 140 kVp) with a SOMATOM Confidence® RT Pro (Siemens Healthineers, Forchheim, Germany) CT scanner. These SECT scans were all acquired with a constant CT dose index (32 cm) of 40 mGy and were reconstructed with the FBP algorithm (B30 kernel, Siemens terminology) in a 512x512 image matrix with a voxel dimension of 0.7x0.7x1.0 mm³. The specifications of the X-ray tube, the spectral filtration, and the detector energy response curve provided by the manufacturer were used to build the CT simulation model in VOXSI.

1 $Z_{\text{eff}} = \sqrt[\beta]{\frac{\sum_i (\omega_i Z_i^{n+1})/A_i}{\sum_i (\omega_i Z_i)/A_i}}$, where $\beta = 3.3$ and i ranges over the n atomic constituents composing a material and the weights ω_i correspond to the weight percentage of element Z_i with atomic number A_i .

The acquired CT image data of the RMI 467 phantom was loaded in *VOXSI* to build a model of the phantom's central slice using the region growing delineation tool². This model is independent from the CT scanner. First, the phantom bulk and its inserts were delineated, and based on the delineation, *VOXSI* automatically calculates a unique material map phantom. Subsequently, the model of the RMI 467 phantom was simulated for the same three X-ray tube potentials and reconstructed with the FBP algorithm in the same matrix size and voxel dimensions as the acquired CT scans. The noise levels measured in the RMI 467 phantom bulk of the real CT scans were adopted in the *VOXSI* CT simulations of the RMI 467 phantom by adding realistic Poisson noise in the projection domain by giving an mAs value ([Supplementary Material B](#)). A visual comparison between a simulated and a measured RMI 467 phantom at 120 kVp is added to [Supplementary Material B](#).

Circular regions of interest within the tissue mimicking inserts were defined on the acquired and the simulated CT scans to calculate the mean Hounsfield Unit (HU) values and their standard deviations for every tissue mimicking insert and every X-ray tube potential.

XCAT phantom dual-energy CT simulations

VOXSI's RMI 467 phantom model was used to perform 80 and 140 kVp DECT simulations. Because the X-ray tube current (mAs) is not equally divided over the two energies in a clinical DECT scanner, and must be known by the simulation algorithm to determine the noise level, an extra dual-spiral DECT scan (80 and 140 kVp) with a total CT dose index (32 cm) of 40 mGy was acquired. The RMI 467 phantom DECT simulations were then used to acquire the RED and Z_{eff} calibration fits ([Figure 3](#), left panels).

An XCAT phantom of the abdominal region was generated with a voxel dimension of $0.7 \times 0.7 \times 1.0 \text{ mm}^3$. Eight different ICRU tissues, from *VOXSI*'s customizable material database, were assigned in XCAT's voxel space [40]. 80 and 140 kVp CT imaging protocols, with FBP reconstructions, of the XCAT phantom were then simulated with the known scanner characteristics.

2 A model of the RMI 467 phantom is provided in the software as standard.

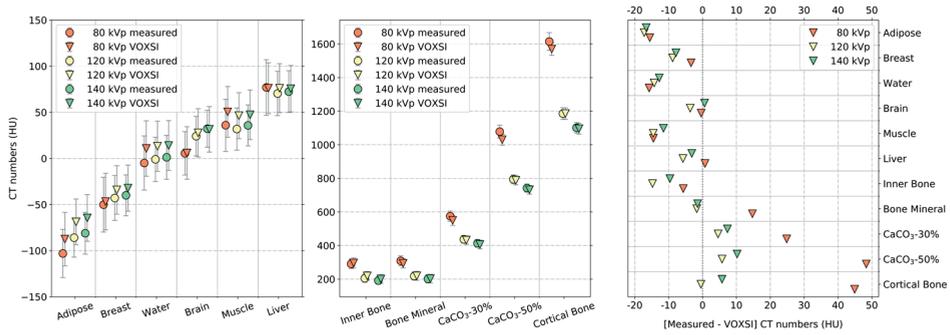


Figure 2. Comparison of the measured and simulated CT numbers (in Hounsfield Units) averaged in circular regions of interest (± 1 standard deviation (SD)) of the RMI 467 soft tissue mimicking inserts (left panel) and the osseous tissue mimicking inserts (middle panel). The right panel plots the absolute differences between the measured and simulated CT numbers for all RMI 467 phantom inserts and the polychromatic X-ray spectra at 80, 120, and 140 kVp.

Results

Quantitative CT validation: RMI 467 phantom

The overall difference between the simulated and the measured mean HUs were below 17 HU for all soft tissue mimicking inserts and below 48 HU for all osseous tissue mimicking inserts (Figure 2). The maximum HU differences between the measured and the simulated CT numbers were obtained at the lowest X-ray tube potential of 80 kVp and in the densest osseous tissue mimicking inserts such as ‘CaCO₃-50%’ (1.53 g/cm³) and ‘Cortical Bone’ (1.82 g/cm³). Increasing the X-ray tube potential to 120 kVp and 140 kVp reduced this maximum error in osseous tissue mimicking inserts to 10 HU and 6 HU respectively.

Dual-energy CT simulations: XCAT phantom

Large RED errors occur around the body and near tissue transitions with a large density change. In the soft tissues, the RED error is smaller ($\pm 4\%$) but increases in regions subject to CT artifacts, i.e. beam hardening. The Z_{eff} errors have a large variation in the soft tissue range ($\pm 12\%$), but a smaller variation in the osseous tissues ($\pm 6\%$). A consistent RED and Z_{eff} overestimation of the lung tissue is visible in Figure 3 (right panels).

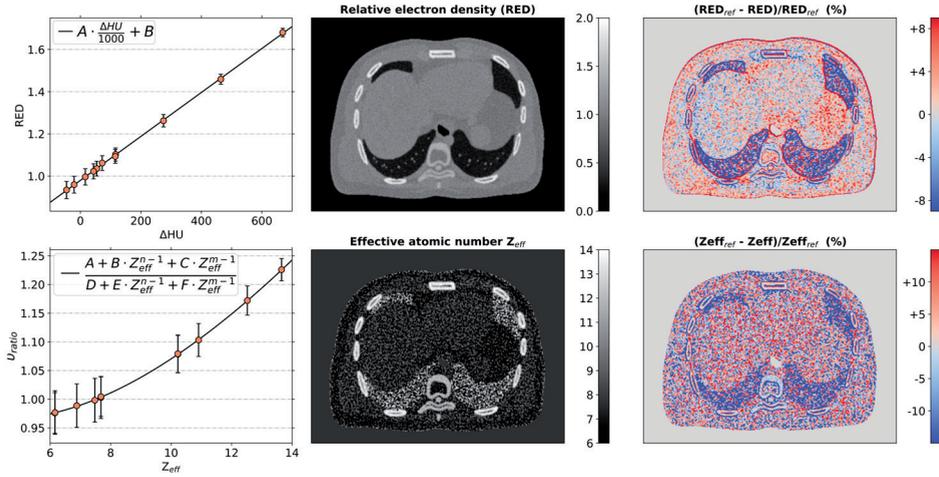


Figure 3. The relative electron density (RED) and effective atomic number (Zeff) calibration curves fitted on the simulated image data of the RMI 467 phantom (left panels) (± 1 SD). In the Zeff calibration curve, only 8 of the 11 data points are clearly visible because 3 points closely overlap with another data point. The RED and Zeff images of the simulated XCAT phantom are depicted in the middle panels. The relative differences between the calculated images and the ground truth are shown in the right panels. (Abbreviations: SD, standard deviation)

Table 1. The coefficients obtained for the relative electron density (RED) and effective atomic number (Zeff) fit. The average residuals for the RED and the Zeff fit were 0.01 and 5.38E-4, respectively.

Fit	A	B	C	D	E	F	m	n	α	R ²
Z _{eff}	0.19	0.01	2.91E-5	0.19	0.01	2.91E-7	4.55	2.87	-	0.99
RED	1.04	0.98	-	-	-	-	-	-	0.90	0.99

Discussion

An SECT and DECT scenario simulator and image analysis toolkit, was presented for the simulation of voxelized volumes of interest in radiology and radiotherapy. The software, which is released for general research use, provides multiple degrees of freedom in the X-ray spectra, geometry, geometrical setup, detector energy response, reconstruction algorithms, and the image analysis. The simulation algorithm accuracy was quantitatively evaluated against a RMI 467 phantom with multiple tissue mimicking inserts that was scanned with a third generation CT scanner at 80, 120, and 140 kVp. The versatility of the software was demonstrated by an elaborated research example investigating the accuracy of the RED and Z_{eff} determination algorithms in a 3D humanoid thoracic XCAT phantom with known tissue characteristics.

Based on the results presented in quantitative CT evaluation (Figure 2), we expect that the relatively small differences (± 15 HU) in CT numbers were not only originating from the accuracy of the CT simulation algorithm and the reconstruction algorithm, but also from the accuracy of the reported atomic compositions of the tissue mimicking inserts. The uncertainties of the atomic compositions of the RMI 467 phantom are not reported in their certificate. Additionally, the small differences in CT numbers could also be explained by the fact that the beam hardening correction algorithm implemented in the software is different from the undisclosed beam hardening correction algorithm in the commercial CT reconstruction software. Eventually, we expect the two different beam hardening correction implementations to be the main cause of discrepancies between the measured and the simulated CT numbers, mainly for the water insert where the Hounsfield Units should be around 0 HU, according to Hounsfield's definition. However, for radiotherapy treatment planning with MV photon beams, 4% to 10% uncertainty in RED would result in no more than a 2% error in dose [41-44].

In Figure 3, the maximum RED errors occur around the body contour and near the regions with large density changes ($\pm 12\%$): e.g. trachea, lung, and bones, because of the boundary smoothing caused by the FBP image reconstruction algorithm. The Z_{eff} calibration curve shows that a small change in the v_{ratio} will result in a large shift in Z_{eff} value. This gives rise to noisy Z_{eff} images, and therefore larger errors between the calculated Z_{eff} image and the ground truth. The largest relative difference in Z_{eff} is found in lung tissue because the Z_{eff} of ICRU inflated lung tissue has a very low RED but a normal Z_{eff} in the soft tissue range.

The CT simulation software can provide researchers with simulated raw projection data, which is more difficult to obtain from clinical CT scanners, to investigate e.g. material decomposition algorithm physical quantity extraction, SECT RED calculation, or projection based (metal) artifact correction algorithms, to name just a few applications [45-47]. Currently, the ASTRA toolbox and the MATLAB implementation of FBP were implemented in the software. As an alternative, the open-source reconstruction toolkit (RTK) can be used to reconstruct the simulated projections [48]. RTK is mostly used to reconstruct cone-beam CT projections, but can also be used to reconstruct plane detector fan-beam CT projections, where no rebinning of fan projections to parallel projections is required. The ASTRA reconstructions in the section: “Quantitative CT validation” were repeated with RTK, no noteworthy differences in CT numbers were found.

The image analysis toolkit in the software provides tools to calculate image metrics for the SECT scans, and tools to perform a DECT analysis. A potential use of the CT scenario generator is to find optimal X-ray spectra, filtration materials or thicknesses for a DECT-based tissue assignment in patient-like geometries, given a user specified geometry (e.g. XCAT phantom) and detector energy response, which is use-

ful for MC dose calculations in megavoltage photon, proton, and brachytherapy treatments. Different tissue assignment methods could be compared directly because the software still has the originally assigned materials in voxel space. In a real patient the absolute truth is missing and therefore only relative differences can be shown between various SECT or DECT based tissue assignment methods, e.g. comparisons of radiotherapy dose distributions [16, 49].

The presented software can also assess different RED or Z_{eff} quantification algorithms and tissue assignment methodologies. Currently, the latter are based on individual pixel values: low and high energy CT numbers, RED, and Z_{eff} , which have three major disadvantages. First, surrounding pixel values or body site information are not taken into account. Second, a material with comparable tissue characteristic could be wrongly assigned due to an increased noise level and third, no knowledge of the anatomy is considered. Using the flexibility of the XCAT phantom and other voxelized geometries, alternative methods making use of artificial intelligence e.g. deep learning algorithms, could be developed to improve tissue segmentation considering 3D voxel space and CT artifacts. The user is not limited to the analysis algorithms provided in the GUI. The simulation data can be exported as MATLAB or DICOM files which can be further analyzed in external software packages.

References

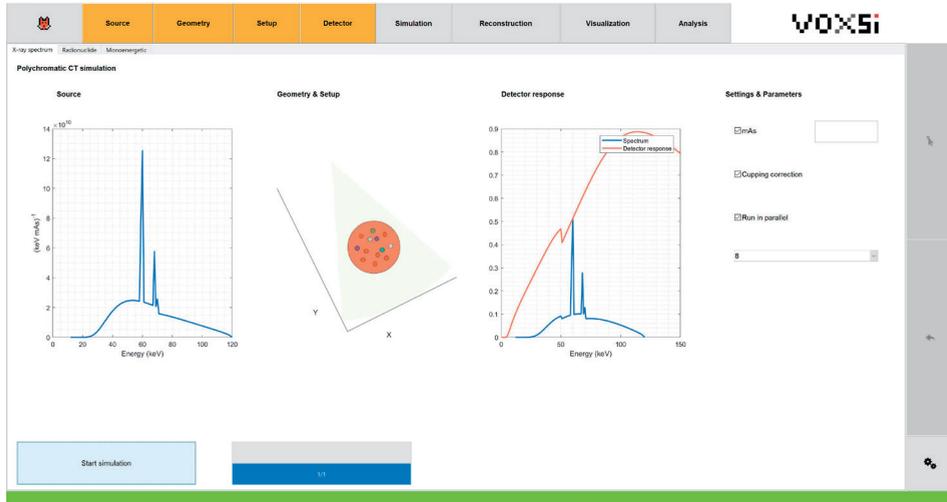
1. Grau, C., et al., *Radiotherapy equipment and departments in the European countries: final results from the ESTRO-HERO survey*. *Radiother Oncol*, 2014. **112**(2): p. 155-64.
2. Chen, G.-P., et al., *Improving CT quality with optimized image parameters for radiation treatment planning and delivery guidance*. *Physics and Imaging in Radiation Oncology*, 2017. **4**: p. 6-11.
3. Hunemohr, N., et al., *Tissue decomposition from dual energy CT data for MC based dose calculation in particle therapy*. *Med Phys*, 2014. **41**(6): p. 061714.
4. Kuchenbecker, S., et al., *Dual energy CT: how well can pseudo-monochromatic imaging reduce metal artifacts?* *Med Phys*, 2015. **42**(2): p. 1023-36.
5. Landry, G., et al., *Sensitivity of low energy brachytherapy Monte Carlo dose calculations to uncertainties in human tissue composition*. *Med Phys*, 2010. **37**(10): p. 5188-98.
6. Malusek, A., et al., *The potential of dual-energy computed tomography for quantitative decomposition of soft tissues to water, protein and lipid in brachytherapy*. *Phys Med Biol*, 2013. **58**(4): p. 771-85.
7. Tawfik, A.M., et al., *Dual-energy CT of head and neck cancer: average weighting of low- and high-voltage acquisitions to improve lesion delineation and image quality-initial clinical experience*. *Invest Radiol*, 2012. **47**(5): p. 306-11.
8. Vaniqui, A., et al., *The impact of dual energy CT imaging on dose calculations for pre-clinical studies*. *Radiat Oncol*, 2017. **12**(1): p. 181.
9. Wichmann, J.L., et al., *Virtual monoenergetic dual-energy computed tomography: optimization of kiloelectron volt settings in head and neck cancer*. *Invest Radiol*, 2014. **49**(11): p. 735-41.
10. Yang, M., et al., *Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues*. *Phys Med Biol*, 2010. **55**(5): p. 1343-62.
11. Zhou, C., et al., *Monoenergetic imaging of dual-energy CT reduces artifacts from implanted metal orthopedic devices in patients with fractures*. *Acad Radiol*, 2011. **18**(10): p. 1252-7.
12. Bazalova, M., et al., *Dual-energy CT-based material extraction for tissue segmentation in Monte Carlo dose calculations*. *Phys Med Biol*, 2008. **53**(9): p. 2439-56.
13. Landry, G., et al., *Improved dose calculation accuracy for low energy brachytherapy by optimizing dual energy CT imaging protocols for noise reduction using sinogram affirmed iterative reconstruction*. *Z Med Phys*, 2016. **26**(1): p. 75-87.
14. Landry, G., et al., *Extracting atomic numbers and electron densities from a dual source dual energy CT scanner: experiments and a simulation model*. *Radiother Oncol*, 2011. **100**(3): p. 375-9.
15. Vilches-Freixas, G., et al., *Comparison of projection- and image-based methods for proton stopping power estimation using dual energy CT*. *Physics and Imaging in Radiation Oncology*, 2017. **3**: p. 28-36.
16. Wohlfahrt, P., et al., *Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans*. *Int J Radiat Oncol Biol Phys*, 2017. **97**(2): p. 427-434.
17. Hudobivnik, N., et al., *Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images*. *Med Phys*, 2016. **43**(1): p. 495.
18. Gallio, E., et al., *A GPU Simulation Tool for Training and Optimisation in 2D Digital X-ray Imaging*. *PLoS One*, 2015. **10**(11): p. e0141497.

19. Lazos, D., et al., *An integrated research tool for X-ray imaging simulation*. Comput Methods Programs Biomed, 2003. **70**(3): p. 241-51.
20. Son, I.Y., et al., *X-ray imaging optimization using virtual phantoms and computerized observer modelling*. Phys Med Biol, 2006. **51**(17): p. 4289-310.
21. Ay, M.R. and H. Zaidi, *Development and validation of MCNP4C-based Monte Carlo simulator for fan- and cone-beam X-ray CT*. Phys Med Biol, 2005. **50**(20): p. 4863-85.
22. De Man, B., et al., *Metal Streak Artifacts in X-ray Computed Tomography: A Simulation Study*. IEEE, 1999. **46**(3): p. 691-696.
23. Kappler, S., et al., *A full-System simulation chain for computed tomography scanners*. IEEE - Nuclear Science Symposium Conference Record, 2009.
24. Segars, W.P., et al., *Realistic CT simulation using the 4D XCAT phantom*. Med Phys, 2008. **35**(8): p. 3800-8.
25. Segars, W.P., et al., *4D XCAT phantom for multimodality imaging research*. Med Phys, 2010. **37**(9): p. 4902-15.
26. Landry, G., F. deBlois, and F. Verhaegen, *ImaSim, a software tool for basic education of medical X-ray imaging in radiotherapy and radiology*. Frontiers in Physics, 2013. **1**.
27. Poludniowski, G., et al., *SpekCalc: a program to calculate photon spectra from tungsten anode X-ray tubes*. Phys Med Biol, 2009. **54**(19): p. N433-8.
28. Poludniowski, G.G., *Calculation of X-ray spectra emerging from an X-ray tube. Part II. X-ray production and filtration in X-ray targets*. Med Phys, 2007. **34**(6): p. 2175-86.
29. Poludniowski, G.G. and P.M. Evans, *Calculation of X-ray spectra emerging from an X-ray tube. Part I. electron penetration characteristics in X-ray targets*. Med Phys, 2007. **34**(6): p. 2164-74.
30. Segars, W.P., et al., *Development of a 4-D digital mouse phantom for molecular imaging research*. Mol Imaging Biol, 2004. **6**(3): p. 149-59.
31. Jacobs, F., et al., *A fast algorithm to calculate the exact radiological path through a pixel or voxel space*. J. Comput Inf Technol, 1998. **6**(1): p. 89-94.
32. Siddon, R.L., *Fast calculation of the exact radiological path for a three-dimensional CT array*. Med Phys, 1985. **12**(2): p. 252-5.
33. Berger, M., et al. *XCOM: Photon Cross Sections Databas*. 1998; Available from: <https://www.nist.gov/pml/xcom-photon-cross-sections-database>.
34. van Aarle, W., et al., *The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography*. Ultramicroscopy, 2015. **157**: p. 35-47.
35. ICRU, *ICRU report 44 - Tissue substitutes in radiation dosimetry and measurement*. 1989.
36. Landry, G., et al., *Simulation study on potential accuracy gains from dual energy CT tissue segmentation for low-energy brachytherapy Monte Carlo dose calculations*. Phys Med Biol, 2011. **56**(19): p. 6257-78.
37. Landry, G., et al., *Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients*. Phys Med Biol, 2013. **58**(19): p. 6851-66.
38. Saito, M. and S. Sagara, *A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body*. Med Phys, 2017. **44**(6): p. 2293-2303.
39. Yu, L., et al., *Virtual monochromatic imaging in dual-source dual-energy CT: radiation dose and image quality*. Med Phys, 2011. **38**(12): p. 6371-9.
40. Saito, M., *Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship*. Med Phys, 2012. **39**(4): p. 2021-30.
41. Cozzi, L., et al., *Dosimetric impact of computed tomography calibration on a*

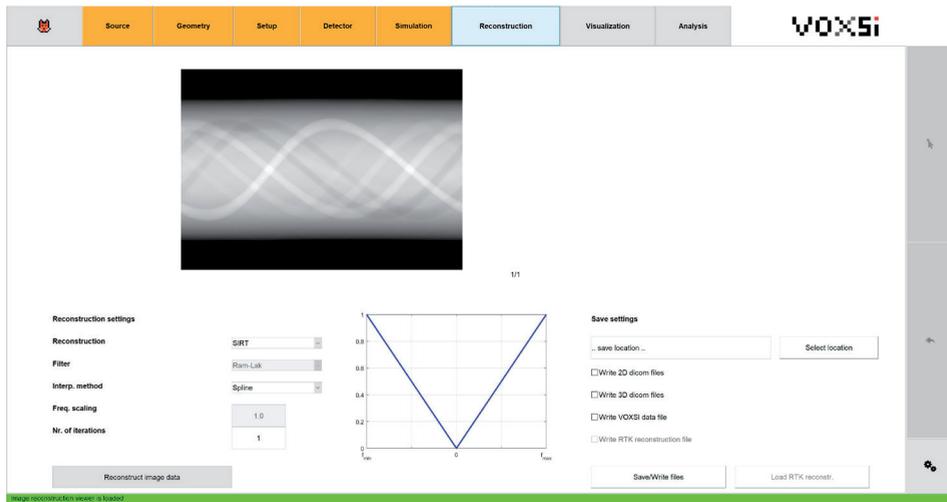
- commercial treatment planning system for external radiation therapy*. Radiother Oncol, 1998. **48**(3): p. 335-8.
42. Ducros, N., et al., *Regularization of non-linear decomposition of spectral X-ray projection images*. Med Phys, 2017. **44**(9): p. e174-e187.
 43. Geise, R.A. and E.C. McCullough, *The use of CT scanners in megavoltage photon-beam therapy planning*. Radiology, 1977. **124**(1): p. 133-41.
 44. Kilby, W., J. Sage, and V. Rabett, *Tolerance levels for quality assurance of electron density values generated from CT in radiotherapy treatment planning*. Phys Med Biol, 2002. **47**(9): p. 1485-92.
 45. Nawaz, S., J. Fu, and D. Fan, *Metal artifacts reduction in X-ray CT based on segmentation and forward-projection*. Biomed Mater Eng, 2014. **24**(6): p. 3287-93.
 46. Prell, D., et al., *A novel forward projection-based metal artifact reduction method for flat-detector computed tomography*. Phys Med Biol, 2009. **54**(21): p. 6575-91.
 47. van der Heyden, B., et al., *Clinical evaluation of a novel CT image reconstruction algorithm for direct dose calculations*. Physics and Imaging in Radiation Oncology, 2017. **2**: p. 11-16.
 48. Rit, S., et al., *The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)*. J Phys Conf Series, 2014.
 49. Wohlfahrt, P., et al., *Dual-energy CT based proton range prediction in head and pelvic tumor patients*. Radiother Oncol, 2017. **125**(3): p. 526-533.

Supplementary Material

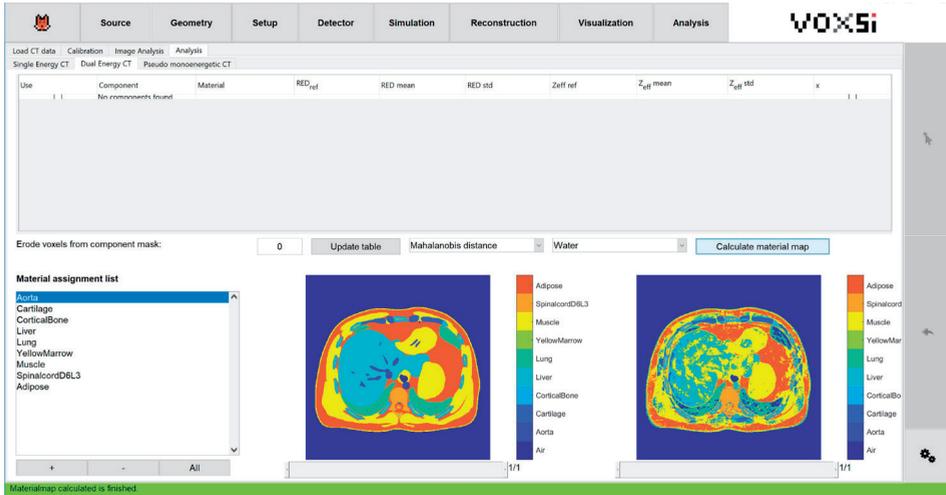
A. VOXSI interface snapshots



Supplementary 1. CT simulation viewer, after selection of the 120 kVp source, the RMI 467 phantom geometry, the CT setup, and the detector energy response.

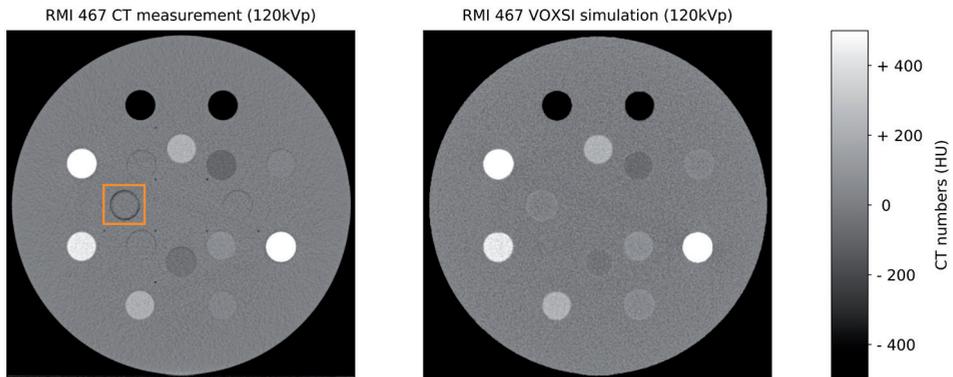


Supplementary 2. CT projection and reconstruction viewer, wherein the raw simulated projection data is visualized. In this example, the projection data of the RMI 467 tissue characterization phantom is shown in the viewer.



Supplementary 3. Dual-energy CT (DECT) analysis toolkit after the tissue assignment using the Mahalanobis distance, based on the relative electron density (RED) and the effective atomic number (Z_{eff}), with materials from a material assignment list which can be adjusted according to the user requirements.

B. CT measurement and VOXSI CT simulation: a visual comparison



Supplementary 4. The CT measurement of the RMI 467 tissue characterization phantom (left panel) compared to the *VOXSI* simulation (right panel). The surrounding plastic tube of the liquid water insert (inside the orange square) was not modelled in *VOXSI*. In the CT measurement, three solid water inserts (bulk material) were used to fill empty insert positions and were only visible due to the submillimeter air gap between the inserts and the phantom bulk. These solid water inserts were not simulated in *VOXSI*.

3

Chapter 3

Clinical evaluation of a novel CT image reconstruction algorithm for direct dose calculations: DirectDensity™

Brent van der Heyden, Michel Öllers, André Ritter, Frank Verhaegen, and
Wouter van Elmpt.

Published in

Physics and Imaging in Radiation Oncology (2017)
<http://dx.doi.org/10.1016/j.phro.2017.03.001>

Abstract

Background and Purpose: Computed tomography (CT) imaging is frequently used in radiation oncology to calculate radiation dose distributions. In order to calculate doses, the CT numbers must be converted into densities by an energy dependent conversion curve. A recently developed algorithm directly reconstructs CT projection data into relative electron densities which eliminates the use of separate conversion curves for different X-ray tube potentials. Our work evaluates this algorithm for various cancer sites and shows its applicability in a clinical workflow.

Materials and Methods: The RMI 467 tissue characterization phantom (Gammex) was scanned to characterize the CT number to density conversion curves. In total, 33 patients with various cancer sites were scanned using multiple tube potentials. All CT acquisitions were reconstructed with the standard Filtered Back-Projection (FBP) and the new developed DirectDensity™ (DD) algorithm (Siemens Healthineers). The mean tumor doses and the volume percentage that receives more than 95% of the prescribed dose were calculated for the planning target volume. Relevant parameters for the organs-at-risk for each tumor site were also calculated.

Results: The relative mean dose differences between the standard 120 kVp FBP CT scan workflow and the DD CT scans (80, 100, 120 and 140 kVp) were in general less than 1% for the planned target volume and organs-at-risk.

Conclusion: The energy independent DD algorithm allows for accurate dose calculations over a variety of body sites. This novel algorithm eliminates the tube potential specific calibration procedure and thereby simplifies the clinical radiotherapy workflow.

Conflict of Interest

This study was partly supported by Siemens Healthineers (Forchheim, Germany).

Acknowledgements

Vittorio Colombo from Siemens Healthineers is acknowledged for his valuable comments during this work.

Introduction

Computed Tomography (CT) imaging is the most frequently used modality for radiotherapy treatment planning, i.e. organ delineation and dose calculations [1, 2]. In most cases, a fixed X-ray tube potential (e.g. 120 kVp) is chosen during the CT acquisition, although different X-ray tube potentials may offer better contrast or signal-to-noise ratio for individual patients [3-5].

The CT scan of the patient is an essential part of the radiation dose calculation procedure in the treatment planning system (TPS). After delineation of target volumes and organs-at-risk (OAR) the treatment plan is evaluated based on the dose constraints of the OAR. Before a dose distribution can be calculated in the TPS, the CT numbers must be converted into relative electron or mass densities. Depending on the dose calculation method, either relative electron density (RED) or mass density (MD) is used. The CT number to density conversion is performed using a calibration curve based on phantom measurements with inserts of known RED and MD [6-8]. The shape and characteristics of the conversion curve is dependent on the chosen X-ray tube potential (kVp) during the calibration process.

A novel image reconstruction algorithm was recently developed that directly reconstructs raw CT projection data into RED images. In contrast to conventional CT images, where the CT numbers quantify the linear photon attenuation and therefore depends on the energy of the X-ray photons, the novel image reconstruction algorithm quantifies the RED of the scanned object which does not depend on the X-ray photon energy, but only on the composition of the scanned object itself.

This image reconstruction algorithm eliminates the use of a CT number to density conversion curve for each selectable X-ray tube potential in the TPS and instead a single curve can be used. Therefore, this image reconstruction algorithm has the potential to make the treatment planning workflow simpler and more robust.

Our study evaluates the accuracy of this image reconstruction algorithm in radiotherapy treatment plans by quantifying differences between the dose distributions. The differences are calculated for different cancer sites between the dose distributions of the conventional 120 kVp CT images and the RED CT images obtained with different X-ray tube potentials.

Materials and Methods

DirectDensity™ image reconstruction algorithm

The developed image reconstruction algorithm (DirectDensity™, Siemens Healthcare GmbH, Germany) operates in both the projection and the image space whereby raw single-energy CT projection data is reconstructed into CT images which are

interpretable as RED images, independent from the selected X-ray tube potential. The DD image reconstruction algorithm achieves a material decomposition into water and bone, based on a threshold value for bone. Effective water and bone thicknesses are obtained by forward projecting the CT images after bone thresholding. The final RED line integrals are calculated by summing the multiplication values of each material effective thickness and its corresponding RED. A more detailed explanation of the DD image reconstruction algorithm is given in [Supplemental Material A](#).

A compact notation was introduced here to denote the combinational use of the image reconstruction algorithm and the X-ray tube potential. In this notation, the first part indicates the image reconstruction algorithm: ‘FPB’ is used for filtered back-projection (Siemens B30 kernel) and ‘DD’ is used for DirectDensity™ (Siemens E30 kernel). The following number in the notation indicates the X-ray tube potential, e.g. DD-140kVp defines a CT scan acquired at an X-ray tube potential of 140 kVp and reconstructed with the novel image reconstruction algorithm.

Phantom calibration

Both the FBP and the DD image reconstruction algorithms were validated with the RMI 467 tissue characterization phantom (Gammex, Middleton, WI). This phantom with tissue mimicking inserts (e.g. lung, adipose, soft tissues, calcium and iodine) was scanned at three different X-ray tube potentials (80, 120 and 140 kVp) with a SOMATOM Confidence® RT Pro (Siemens Healthineers, Forchheim, Germany) 64-slice CT scanner. The three CT scans of the phantom were obtained with a tube current of 250 mAs and reconstructed with a 3 mm slice thickness.

To obtain CT number to RED or MD curves, all three CT scans of the RMI 467 phantom were reconstructed with the DD algorithm, and the 120kVp CT scan was additionally reconstructed with the FBP algorithm.

The data points of the CT number to density conversion curves were obtained by calculating mean CT numbers in a region of interest with respect to the certified RED and MD of the tissue mimicking inserts. The unit of the CT numbers depends on the image reconstruction algorithm: the FBP algorithm reconstructs the raw projection data in Hounsfield Units (HU) and the DD algorithm provides a simple CT number scaling proportional to the RED, described by [Equation 1](#).

Further in this study, the term ‘CT number’ is only used as a general term to describe pixel values of a CT image reconstruction. In order to maintain the consistency between the unit of the CT numbers, CT_{HU} is used to denote an FBP image reconstruction in HU, CT_{RED} is used to denote a DD image reconstruction in the linear RED scaling (Eq. 1).

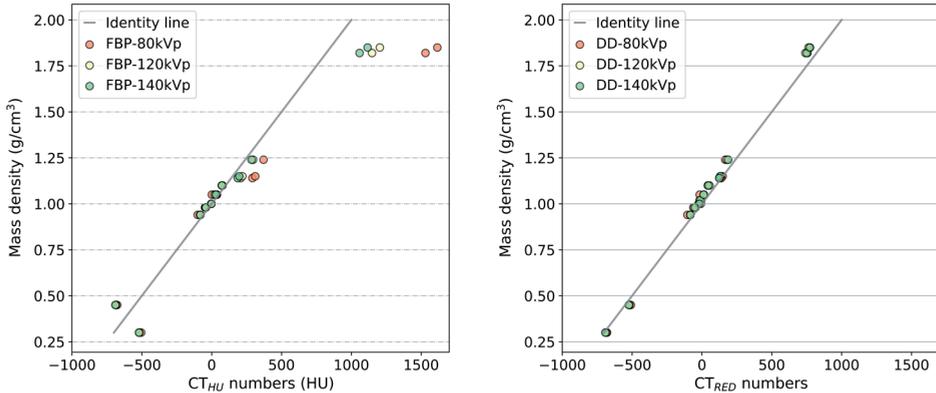


Figure 1. The CT_{HU} number to mass density (MD) conversion curves (left panel) obtained when the CT scans of the RMI 467 phantom and its tissue mimicking inserts were reconstructed with the filtered back-projection (FBP) image reconstruction algorithm and the energy independent CT_{RED} number to MD conversion curve (right panel) obtained when the same CT scans were reconstructed with the DirectDensity™ (DD) image reconstruction algorithm.

$$RED = \frac{CT_{RED}}{1000} + 1 \quad (\text{Eq. 1})$$

To prevent negative RED values in Equation 1, CT_{RED} numbers between -1024 and -1000 were forced to be -1000 by applying thresholding. The mean CT_{RED} numbers obtained from the DD-80kVp, DD-120kVp and DD-140kVp image reconstructions were combined into a mean data point of the energy independent CT_{RED} number to density conversion curve for the novel image reconstruction algorithm. In this fit we assumed the energy independence of the CT_{RED} numbers. The CT_{HU} number to density conversion curve for the FBP-120kVp image reconstruction was obtained by calculating mean HU in a region of interest as a function of the phantom insert densities. The CT number to density conversion curve in the TPS was interpreted as a linear interpolation between the mean data points and was not used as a linear fit between all data points.

The CT_{RED} number to density conversion curve, shown in the right panel of Figure 1, was used in the TPS to perform dose calculation on the DD image reconstructions obtained at different X-ray tube potentials.

Patient population

The DD image reconstruction algorithm was evaluated in a clinical imaging and radiation treatment workflow. A group of 26 patients with various cancer sites were scanned with two different imaging protocols (Table 1). First the patient was scanned with the conventional X-ray tube potential of 120 kVp, whereafter an additional

dual-spiral dual-energy (DE) CT acquisition was made at 80 kVp and 140 kVp. The CT scans were performed according to a clinical imaging protocol whereby an equivalent imaging dose was opted between the first 120 kVp CT acquisition and the second dual-spiral DECT acquisition.

The CT scans were reconstructed to obtain four image datasets for each patient: DD-80kVp, DD-120kVp, DD-140kVp and FBP-120kVp. Patients were delineated and planned (Eclipse™ version 11, Varian Medical Systems, Palo Alto, CA) based on the FBP-120kVp image reconstruction. Both the clinical contours and the clinical treatment plan were directly copied to the DD-80kVp, DD-120kVp and DD-140kVp images. Due to the sequential acquisitions at different X-ray tube potentials, patient motion could take place between the three CT acquisitions. No correction was applied to compensate for these deformations.

Additionally, a cohort of seven breast cancer patients was used to investigate the automatic adjustment of the X-ray tube potential and the tube current based on the topogram [9] (CARE kV, Siemens Healthcare GmbH, Germany). The standard CT scan was performed with a conventional X-ray tube potential of 120 kVp and if a different tube potential was suggested (mainly 100 kVp) also this image was acquired. An Institutional Review Board approval was given for all patients included in our study.

Dose distribution analysis

Dose calculations, calculated with Acuros® XB (v.11.0.31, Eclipse, Varian Medical Systems, Palo Alto, CA) and reported in dose to medium, were performed on all CT image reconstructions of each patient included in our study [9-13]. The Acuros® dose calculation engine uses MD to assign materials inside the voxels, hence a CT number to MD conversion curve was required. Note that a CT number to RED conversion must be used if the dose calculation engine scales the pre-calculated dose kernels according to the RED matrix as is done e.g. in the analytical anisotropic algorithm (AAA, Eclipse™, Varian, Palo Alto, CA) [14].

Analysis and statistics

The first group of 26 patients was subdivided into five smaller patient subgroups, the group of seven breast cancer patients was considered as the sixth subgroup.

The subgroups and the number of patients inside each subgroup, indicated between brackets, were subsequently listed: prostate cancer (N=3), rectal cancer (N=4), bone metastasis (N=7), brain metastases (N=5), breast cancer (N=7) and miscellaneous tumor sites (N=7). The miscellaneous tumor site group consisted of cancers in the bladder, spleen, stomach, groin metastasis and seminoma testis metastasis patients.

For each of the six subgroups, mean doses and the volume percentage that receives more than 95% of the prescribed dose ($V_{95\%}$) were calculated in the planning target volume (PTV). Because the OAR delineation depended on the tumor position, the treatment plan and intent, the same OAR were not always delineated throughout the entire cancer patient subgroup. The mean doses delivered in the OAR were only calculated for the prostate, rectum and lung cancer patient subgroup. For each of the OAR, the dose volume histogram parameters that were used as planning constraints were compared between the different dose distributions.

Results

Calibration and phantom evaluation

The left panel of **Figure 1** shows the energy dependency of the HU on the material inserts. The energy dependence was reduced using the DD algorithm as shown in the conversion curve for the DD-80kVp, DD-120kVp and DD-140kVp image reconstructions (**Figure 1**, right panel). The CT_{RED} numbers which are proportional to the RED (Eq. 1) had a $-0.1\% \pm 2.2\%$ (1SD) residual from the linear identity line for all tissue mimicking inserts and all X-ray tube potentials. Individually, relative differences up to $-0.1\% \pm 1.8\%$ were observed for the DD-120kVp image reconstruction, $-0.1\% \pm 2.8\%$ for the DD-80kVp image reconstruction and $0.0\% \pm 1.9\%$ for the DD-140kVp image reconstruction.

Patient dose distributions

Relative differences in mean dose and in volumetric dose coverage were calculated between the clinically used FBP-120kVp image reconstruction and the other DD-80kVp, DD-100kVp (breast cancer patient subgroup only), DD-120kVp and DD-140kVp image reconstructions. The relative mean dose differences and the mean differences in the $V_{95\%}$ were averaged over all cancer patients in a specific subgroup and are listed in **Table 1** and **Table 2**. All relative differences and standard deviations in mean dose and $V_{95\%}$ for each treatment site were smaller than 1%, except for the dose differences in the anal canal. Here, the largest relative mean dose differences of almost 4% and large standard deviations of 2.3% were found in the DD-80kVp and 140kVp image reconstruction and are most likely caused by the small organ displacement in the images. The relative mean dose differences and standard deviation of the anal canal between the DD-120kVp and DD-140kVp image reconstructions were both less than 1%.

Table 1. Relative mean dose differences ($\Delta \bar{D}$), differences in the volume percentage that receives more than 95% of the prescribed dose ($\Delta V_{95\%}$) and standard deviations are calculated in the planning target volume (PTV) or organs-at-risk between two image reconstructions per cancer patient subgroup. The total number of cancer patients in the subgroup (N) is indicated in brackets.

Relative difference	FBP-120kVp DD-80kVp (%)	FPB-120kVp DD-120kVp (%)	FBP-120kVp DD-140kVp (%)
Prostate (N=3)			
$\Delta \bar{D}_{PTV}$	0.0±0.6	0.4±0.6	0.4±0.5
$\Delta V_{95\%,PTV}$	-0.9±0.1	0.7±0.9	0.2±0.5
$\Delta \bar{D}_{Rectum}$	-0.2±0.9	0.7±0.6	0.1±0.9
$\Delta \bar{D}_{Anal\ canal}$	3.5±2.2	0.7±0.5	3.7±2.3
$\Delta \bar{D}_{Bladder}$	0.4±0.8	0.4±0.4	0.7±0.7
Rectum (N=4)			
$\Delta \bar{D}_{PTV}$	-0.2±0.4	0.1±0.1	0.0±0.3
$\Delta V_{95\%,PTV}$	-0.2±0.2	0.0±0.1	0.0±0.2
$\Delta \bar{D}_{Bowels}$	0.2±0.7	0.2±0.2	0.2±0.5
$\Delta \bar{D}_{Bladder}$	-0.1±0.3	0.2±0.1	0.0±0.3
Bone metastasis (N=7)			
$\Delta \bar{D}_{PTV}$	-0.1±0.3	0.1±0.1	0.1±0.3
$\Delta V_{95\%,PTV}$	-0.2±0.5	0.1±0.2	0.1±0.4
Head (N=5)			
$\Delta \bar{D}_{PTV}$	-0.4±0.4	-0.3±0.5	-0.4±0.4
$\Delta V_{95\%,PTV}$	0.5±1.0	0.4±0.7	0.5±1.2
Miscellaneous (N=7)			
$\Delta \bar{D}_{PTV}$	-0.2±0.4	-0.0±0.3	-0.2±0.6
$\Delta V_{95\%,PTV}$	-0.4±0.7	-0.3±0.8	-0.1±0.9

Table 2. Relative mean dose differences ($\Delta \bar{D}$), differences in the volume percentage that receives more than 95% of the prescribed dose ($V_{95\%}$) and standard deviations are calculated in the planning target volume (PTV) or organs-at-risk between the FBP-120kVp and DD-100kVp image reconstructions of the 7 breast cancer patient analyzed.

Relative difference Breast (N=7)	FBP-120kVp DD-100kVp (%)
$\Delta \bar{D}_{PTV}$	-0.3±0.2
$\Delta V_{95\%,PTV}$	-0.6±0.8
$\Delta \bar{D}_{Heart}$	0.2±0.4
$\Delta \bar{D}_{Left\ lung}$	0.3±0.4
$\Delta \bar{D}_{Right\ lung}$	0.0±0.2

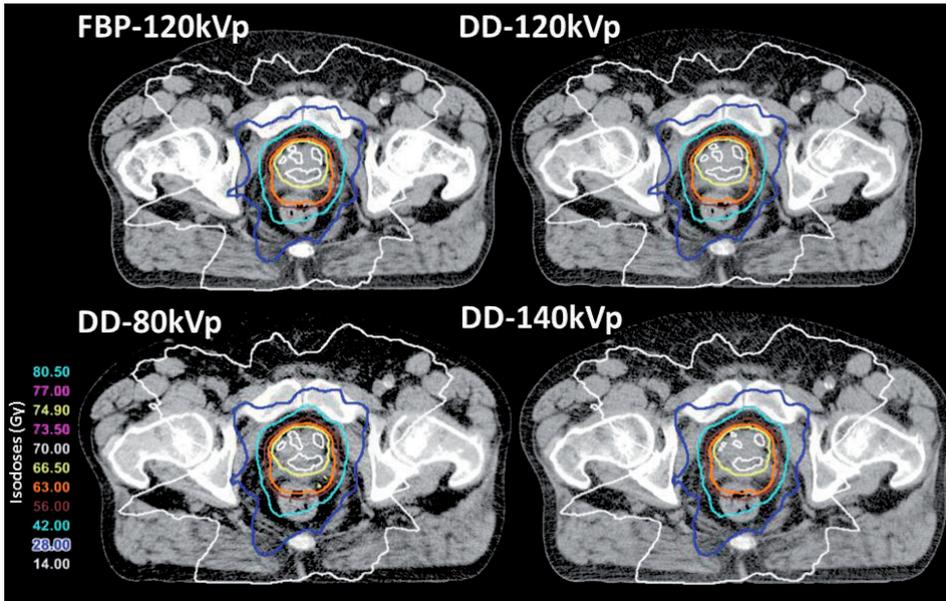


Figure 2. Dose distributions of a clinical prostate cancer treatment plan calculated on the four different image reconstructions (window width/level = 300/40 HU) ($CTDI_{vol,80kVp} = 7.9$ mGy; $CTDI_{vol,140kVp} = 8.7$ mGy; $CTDI_{vol,120kVp} = 17.7$ mGy).

Figure 2 and Figure 3 show the isodose lines of a prostate and a head-and-neck cancer patient, respectively. These isodose lines of both patients showed similar patterns between the dose distributions calculated based on different image reconstructions. Both patients had the largest relative mean dose differences in their patient group, due to their large body volume more voxels could be assigned incorrectly in the density to material conversion which is done before the dose calculation is started, although the relative difference for both tumor and OAR were smaller than 1%. The isodose lines for other cancer sites showed similar results.

The higher image noise in the dual-spiral DECT images compared to the 120 kVp CT images was due to the lower volume CT dose index ($CTDI_{vol}$) of the separate 80 kVp and 140 kVp CT acquisitions compared to the 120 kVp CT acquisition. A lower imaging dose results in a CT image with more noise. The total imaging dose of the 120 kVp CT scan is divided over the 80 kVp and the 140 kVp CT acquisitions. The CT images obtained through the DD image reconstruction algorithm gave a better image contrast inside dense regions (e.g. bone) compared to the standard FBP image reconstruction.

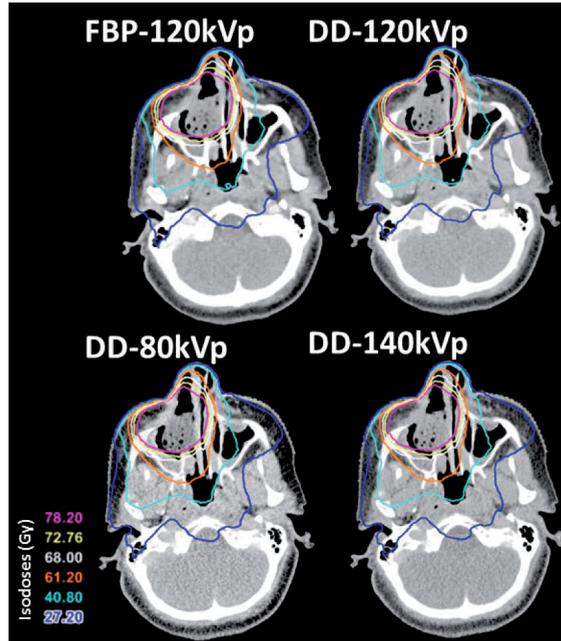


Figure 3. Dose distributions of a clinical treatment plan of a head and neck cancer patient recalculated on the four different image reconstructions (window width/level = 300/ 40 HU) ($CTDI_{vol,80kVp} = 1.3$ mGy; $CTDI_{vol,140kVp} = 2.0$ mGy; $CTDI_{vol,120kVp} = 3.3$ mGy).

Discussion

We evaluated a novel method to simplify the clinical radiation therapy workflow. The CT_{RED} number to MD conversion curve needed for the DD image reconstruction algorithm was obtained through a calibration with the RMI 467 phantom and its certified tissue mimicking inserts and was independent of tube potential. The calibrated CT_{RED} number to MD conversion curve (Figure 1, right panel) was implemented in the TPS and could be used to perform dose calculations on DD image reconstructions independent from the X-ray tube potential of the CT scanner. The DD image reconstruction algorithm used in our study is based on the FBP image reconstruction algorithm, as also the conventional CT images in this study. We evaluated a novel method to simplify the clinical radiation therapy workflow. The CT_{RED} number to MD conversion curve needed for the DD image reconstruction algorithm was obtained through a calibration with the RMI 467 phantom and its certified tissue mimicking inserts and was independent of tube potential.

The calibrated CT_{RED} number to MD conversion curve (Figure 1, right panel) was implemented in the TPS and could be used to perform dose calculations on DD

image reconstructions independent from the X-ray tube potential of the CT scanner. The DD image reconstruction algorithm used in our study is based on the FBP image reconstruction algorithm, as also the conventional CT images in this study. However, there is also the possibility to reconstruct the DD images with an iterative reconstruction (SAFIRE or ADMIRE, Siemens terminology) variant of this algorithm (Siemens F30 kernel). This could potentially reduce the imaging dose and noise further [4, 15-18].

Alternatively, a different method that still needs to be investigated could be found to derive the conversion curve: Beaulieu *et al.* [19] used real human tissue compositions, taken from ICRP23 [20] and listed in Schneider *et al.* [8], to fit a linear relationship between MD and RED. This linear fit is then derived from real human tissue compositions instead of tissue mimicking plastics and could possibly be used as a conversion curve of the DD image reconstruction algorithm, although scanner specific changes will not be taken into account.

The relative mean dose differences in the PTV and the OAR were evaluated for all 33 patients with various cancer sites. For each patient, the relative mean dose difference between both image reconstruction algorithms at different X-ray tube potentials is smaller than 1%, excluding the anal canal because of the displacement of this relatively small structure. A dual-spiral DECT scan was used which could cause image reconstructions that were slightly different due to the patient's respiratory motion. When small delineated structures, such as the anal canal (OAR of the prostate cancer patient subgroup), are copied from the FPB-120kVp image reconstruction to other DD image reconstructions obtained from the DECT scan, the anal canal can be slightly displaced. This is supported by the fact that the mean dose difference of the anal canal (prostate cancer patient subgroup) between the FBP-120kVp and the DD-120kVp image reconstructions is much smaller because both DD image reconstructions were performed on the same raw projection data. By changing the clinical imaging workflow towards the DD algorithm, the CT_{RED} number to density conversion curve in the TPS is now independent from the X-ray tube potential of the CT scanner. Only one CT_{RED} number to density conversion curve should be implemented in the TPS to perform dose calculations on CT scans acquired at different kVp settings.

The dose constraint differences between the clinically accepted FBP-120kVp and the DD treatment plans were small, hence the treatment plan based on the novel DD image reconstruction can be seen as clinical acceptable for various X-ray tube potentials. The use of DD reduces both the clinical workload and the risk of selecting incorrect conversion curves in the TPS, but it also offers the opportunity to optimize each CT protocol in terms of image quality without concern for the accuracy of the dose calculations.

Conclusion

An image reconstruction that directly reconstructs relative electron density images independent from the X-ray tube potential was evaluated in a clinical workflow. We found small differences in dose distribution for the PTV and the OAR of various cancer sites. Therefore, this reconstruction algorithm could be advantageous for implementation in the clinic as it has the potential to improve the workflow, selection of different kVp settings and make the treatment planning procedure more robust.

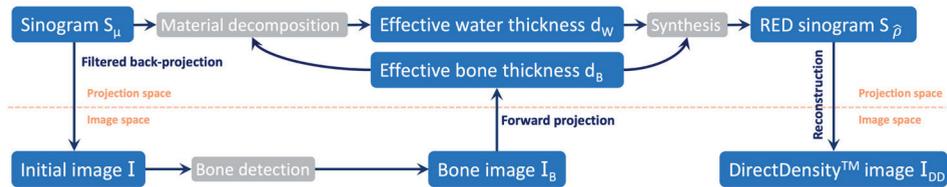
References

1. Grau, C., et al., *Radiotherapy equipment and departments in the European countries: final results from the ESTRO-HERO survey*. *Radiother Oncol*, 2014. **112**(2): p. 155-64.
2. Lu Jiade, J. and W. Brady Luther, in *Radiation oncology an evidence-based approach*. 2008. p. 625-639.
3. Lurz, M., et al., *Automated tube voltage selection in thoracoabdominal computed tomography at high pitch using a third-generation dual-source scanner: image quality and radiation dose performance*. *Invest Radiol*, 2015. **50**(5): p. 352-60.
4. Scholtz, J.E., et al., *Third-generation dual-source CT of the neck using automated tube voltage adaptation in combination with advanced modeled iterative reconstruction: evaluation of image quality and radiation dose*. *Eur Radiol*, 2016. **26**(8): p. 2623-31.
5. Seyal, A.R., et al., *CT of the Abdomen with Reduced Tube Voltage in Adults: A Practical Approach*. *Radiographics*, 2015. **35**(7): p. 1922-39.
6. Constantinou, C., J.C. Harrington, and L.A. DeWerd, *An electron density calibration phantom for CT-based treatment planning computers*. *Med Phys*, 1992. **19**(2): p. 325-7.
7. Das, I.J., et al., *Computed tomography imaging parameters for inhomogeneity correction in radiation treatment planning*. *J Med Phys*, 2016. **41**(1): p. 3-11.
8. Schneider, U., E. Pedroni, and A. Lomax, *The calibration of CT Hounsfield units for radiotherapy treatment planning*. *Phys Med Biol*, 1996. **41**(1): p. 111-24.
9. Fogliata, A., et al., *Dosimetric evaluation of Acuros XB Advanced Dose Calculation algorithm in heterogeneous media*. *Radiat Oncol*, 2011. **6**: p. 82.
10. Kroon, P.S., S. Hol, and M. Essers, *Dosimetric accuracy and clinical quality of Acuros XB and AAA dose calculation algorithm for stereotactic and conventional lung volumetric modulated arc therapy plans*. *Radiat Oncol*, 2013. **8**: p. 149.
11. Onizuka, R., et al., *Accuracy of dose calculation algorithms for virtual heterogeneous phantoms and intensity-modulated radiation therapy in the head and neck*. *Radiol Phys Technol*, 2016. **9**(1): p. 77-87.
12. Tsuruta, Y., et al., *Dosimetric comparison of Acuros XB, AAA, and XVMC in stereotactic body radiotherapy for lung cancer*. *Med Phys*, 2014. **41**(8): p. 081715.
13. Zifodya, J.M., C.H. Challens, and W.L. Hsieh, *From AAA to Acuros XB-clinical implications of selecting either Acuros XB dose-to-water or dose-to-medium*. *Australas Phys Eng Sci Med*, 2016. **39**(2): p. 431-9.

14. Van Esch, A., et al., *Testing of the analytical anisotropic algorithm for photon dose calculation*. Med Phys, 2006. **33**(11): p. 4130-48.
15. Baumueller, S., et al., *Low-dose CT of the lung: potential value of iterative reconstructions*. Eur Radiol, 2012. **22**(12): p. 2597-606.
16. Han, B.K., et al., *Assessment of an iterative reconstruction algorithm (SAFIRE) on image quality in pediatric cardiac CT datasets*. J Cardiovasc Comput Tomogr, 2012. **6**(3): p. 200-4.
17. Rivers-Bowerman, M.D. and J.J. Shankar, *Iterative Reconstruction for Head CT: Effects on Radiation Dose and Image Quality*. Can J Neurol Sci, 2014. **41**(5): p. 620-5.
18. Wortman, J.R., A.J. Adduci, and A.D. Sodickson, *Synergistic Radiation Dose Reduction by Combining Automatic Tube Voltage Selection and Iterative Reconstruction*. J Thorac Imaging, 2016. **31**(2): p. 111-8.
19. Beaulieu, L., et al., *Report of the Task Group 186 on model-based dose calculation methods in brachytherapy beyond the TG-43 formalism: current status and recommendations for clinical implementation*. Med Phys, 2012. **39**(10): p. 6208-36.
20. *International Commission on Radiological Protection, Report on the Task Group on Reference Man*. ICRP Publication No. 23.

Supplementary Material

A. DirectDensity™ image reconstruction algorithm



Supplementary 1. Flow chart of the DirectDensity™ (DD) algorithm.

In **Supplementary 1**, the flow chart of the DirectDensity™ (DD) algorithm is shown. The DirectDensity™ (DD) algorithm starts from the conventional sinogram of a single-energy Computed Tomography (CT) scan. This sinogram contains the values S_μ of the line integrals of the X-ray linear attenuation coefficient μ for each projection ray of the scan. From this sinogram an initial CT image volume with CT numbers I in Hounsfield Units (HU) is reconstructed via filtered back-projection. Effects of beam hardening by water are corrected in this initial image.

In the next step the presence of bone in this volume is detected to calculate a volume image I_B of the isolated bone contribution. For this we make the following assumptions. Volume elements, which have an initial image value below a certain threshold I_T , are classified to contain only a water-equivalent material with a variable density and the elemental composition of pure water. Volume elements, which have an initial image value above this threshold, are classified to contain a mixture of water with a normal density of 1.0 g/cm^3 and a bone-equivalent material. This threshold is determined using calibration phantoms and typically takes values between 50 HU and 300 HU depending on properties of the X-ray spectrum, the filtration and the detector used for CT scanning. The threshold cannot be influenced by the user. We further assume that the contribution of the bone-equivalent material to the total attenuation coefficient in the mixture is proportional to the initial image value. This results in an effective mass density of the bone material within the mixture that is variable and a constant amount of water in the mixture. A suitable elemental composition of the bone material is for example given in du Plessis et al (1998), Med Phys 25, 1195-201. Consequently, the values I_B of the bone image are proportional to the initial value I for initial values $I > I_T$ and zero elsewhere. This bone detection step is most robust if only natural-body materials are present within the volume. The presence of non-natural body materials e.g. metals or contrast agents are not well reflected by the assumptions explained before and can decrease the accuracy of the bone detection. For example, the presence of iodine within tissue can potentially be detected as dense bone due to the higher attenuation coefficient of iodine compared to the basis mate-

rial bone. In the later steps of this algorithm this can lead to an overestimation of electron density.

The resulting bone image I_B is then forward projected. This gives an estimation of the value of line integral s_B of the isolated contribution of the bone-equivalent material within the original CT sinogram. This contribution is also given by the product of the linear attenuation coefficient μ_B and the effective thickness d_B of the bone-equivalent material. Knowing the linear attenuation coefficient μ_B , it is possible to calculate the effective thickness $d_B = \frac{s_B}{\mu_B}$ for each line integral from the contribution s_B .

There is a relationship between the value of the line integral S_μ and the effective thicknesses of bone d_B and water d_W . This relationship is acquired numerically using a simulation that is validated for Siemens CT scanners. The simulation contains a full spectral model of X-rays and their interaction with matter. This relationship is specific to geometric and spectral properties of the CT scan. The specific relationship allows calculating the effective thickness of water $d_W(S_\mu, d_B)$ for each ray as a function of the given value of the line integral S_μ and the previously estimated effective thickness of bone d_B . A complete decomposition into the contributions of the two linear-independent basis materials (i.e. water and bone-equivalent) is achieved in projection space.

In analogy to the conventional CT sinogram, that contains the line integrals of the linear attenuation coefficient, a synthetic sinogram of relative electron density (RED) can be calculated. Knowing REDs $\hat{\rho}_{e,W}$ and $\hat{\rho}_{e,B}$ of water and bone-equivalent material, line integrals $S_{\hat{\rho}_e}$ of the RED for each projection ray are obtained by summing the products of RED and the corresponding effective thickness for each material. The line integrals of RED for all projection rays together form the sinogram of RED.

$$S_{\hat{\rho}_e} = \hat{\rho}_{e,W} \cdot d_W + \hat{\rho}_{e,B} \cdot d_B$$

The final DD images are then reconstructed from the sinogram of RED using either a filtered back-projection or an iterative reconstruction option available at the scanner as requested by the user. The image values I_{DD} reconstructed with the DD algorithm are in an HU-like scaling but proportional to the RED $\hat{\rho}_e$.

$$I_{DD} \approx (\hat{\rho}_e - 1) \cdot 1000$$

The RED of a material is given by the ratio of the material electron density ρ_e and the electron density of water $\rho_{e,W}$.

$$\hat{\rho}_e = \frac{\rho_e}{\rho_{e,W}}$$



Chapter 4

A comparison study between single-and dual-energy CT density extraction methods for neurological proton Monte Carlo treatment planning

Brent van der Heyden*, Isabel P Almeida*, Gloria Vilches-Freixas,
Celine van Beveren, Ana Vaniqui, Carmen Ares, Karin Terhaag,
Gabriel P Fonseca, Daniëlle BP Eekers, and Frank Verhaegen.

* Equal contribution

Published in

Acta Oncologica (2019)

<http://dx.doi.org/10.1080/0284186X.2019.1679879>

Abstract

Background and Purpose: Monte Carlo proton dose calculations requires mass densities calculated from the patient CT image. This work investigates the impact of different single-energy CT (SECT) and dual-energy CT (DECT) to density conversion methods in proton dose distributions for brain tumors.

Materials and Methods: Head CT scans for four patients were acquired in SECT and DECT acquisition modes. Commercial software was used to reconstruct DirectDensity™ images in Relative Electron Densities (RED, ρ_e) and to obtain DECT-based pseudo-monoenergetic images (PMI). PMI and SECT images were converted to RED using piecewise linear interpolations calibrated on a head-sized phantom, these fits were referred to as “PMI2RED” and “CT2RED”. Two DECT-based calibration methods (“Hünemohr-15it” and “Saito-15it”) were also investigated. ρ_e images were converted to mass-densities (ρ_m) to investigate ρ_m differences and one representative patient case was used to make a proton treatment plan. Using CT2RED as reference method, dose distribution differences in the target and in five organs-at-risk (OARs) were quantified.

Results: In the phantom study, Saito-15it and Hünemohr-15it produced the lowest ρ_e root-mean-square error (0.7%) and DirectDensity™ the highest error (2.7%). The proton plan evaluated in the Saito-15it and Hünemohr-15it datasets showed the largest relative differences compared to initial CT2RED plan down to -6% of the prescribed dose. Compared to CT2RED, average range differences were calculated: -0.1 ± 0.3 mm for PMI2RED; -0.8 ± 0.4 mm for Hünemohr-15it, and -1.2 ± 0.4 mm for Saito-15it.

Conclusion: Given the wide choice of available conversion methods, studies investigating the density accuracy for proton dose calculations are necessary. However, there is still a gap between performing accuracy studies in reference ρ_e phantoms and applying these methods in human CT images. For this treatment case, the PMI2RED method was equivalent to the conventional CT2RED method in terms of dose distribution, CTV coverage and OAR sparing, whereas Hünemohr-15it and Saito-15it presented the largest differences.

Acknowledgements

The authors thank Prof. Dr. Katia Parodi, head of the Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München (LMU), for lending the PMMA head size phantom. Dr. Christian Hofmann and Dr. Jens-Christoph Georgi from Siemens Healthineers and Dr. Guillaume Landry from LMU are acknowledged for their valuable comments and discussion on this work.

Introduction

In conventional radiotherapy, single-energy computed tomography (SECT) images are used to perform dose distribution calculations for treatment planning due to their capability of identifying tissue attenuation characteristics. In the treatment planning procedure, sufficient tumor coverage and healthy tissue sparing are desired. The more advanced photon therapies and proton therapy can lead to conformal tumor dose distributions, allowing optimal sparing of organs-at-risk (OARs) [1-3]. With increasing conformity of tumor dose, the accuracy of the dose calculation algorithm and its input from CT imaging becomes more critical. In this study, we will focus on the accuracy of proton dose calculations based on various information extraction methods from CT images.

In a treatment planning system (TPS) for proton radiotherapy the dose calculation engine requests material properties from the patient CT scan, such as stopping power ratios (SPR) or mass densities (ρ_m), depending on the type of dose calculation algorithm. Most proton centers today base their treatment plan on SECT images, although it has been suggested in the literature that dual-energy CT (DECT) can lead to more accurate proton range prediction [4-8]. In proton radiotherapy, various methods are available to convert SECT and DECT images into ρ_e , which can be converted into ρ_m following a simple linear relationship [9].

In radiotherapy clinics, the CT number to density conversion curve is one of the most commonly used methods [10]. This conversion curve is derived from tissue mimicking phantom measurements with known material properties [11, 12]. Alternatively, those tissue mimicking phantom measurement can also be used for the stoichiometric calibration [13] that aims to model CT scanner aspects such as the X-ray spectrum and the detector energy response, and subsequently use this model to predict CT numbers. However, recent studies have shown that stoichiometric methods depend on the tissue mimicking materials used for calibration [14] and concluded that the phantom tissue substitutes used in this work seem to be a good representative of biological tissues and that the “stoichiometric calibration becomes unnecessary and may be replaced by a much simpler SECT calibration procedure [14]”. Hudobivnik *et al.* [5] investigated the accuracy of the stoichiometric calibration and the CT to density conversion curve for proton stopping power ratios calculated using tissue mimicking inserts similar to the insert materials used in this study. They found mean square error differences compared to the ground truth of 1.8% compared to 1.9% respectively. The CT number to density conversion curve and the stoichiometric calibration are both influenced by the X-ray tube potential, so commercial solutions have been developed to provide an energy independent reconstruction algorithm that directly reconstructs CT numbers into relative electron density (ρ_e , RED) units independent from the X-ray tube potential [15].

DECT-based ρ_e conversion methods are less used in clinical routine because radiotherapy treatment planning systems do not provide standardized methods for reading DECT images for dose calculations. In 2016, Wohlfahrt *et al.* [7] discussed the clinical implementation of DECT for proton treatment planning based on pseudo-monoenergetic images (PMIs).

In this study, the ρ_e accuracy of two SECT and three DECT-based ρ_e conversion methods were evaluated in a phantom study with accurately known material properties. Additionally, these five conversion methods were studied in four patient brain CT scans. A proton treatment plan was created for one patient and differences in dose distribution, particularly target coverage and sparing of OARs, were quantified.

Material and Methods

Phantom study and CT protocols

Two ρ_e phantoms with certified chemical compositions and densities were scanned with a SOMATOM Confidence® RT pro 64-slice CT scanner (Siemens Healthineers, Forchheim, Germany) in SECT acquisition mode at 120 kVp and dual-spiral DECT acquisition mode at 80 and 140 kVp. All phantom scans were reconstructed with sinogram affirmed iterative reconstruction (SAFIRE, Siemens terminology) with beam hardening correction for bone enabled (Q34 kernel, strength 3). The SECT scan was also reconstructed with the commercially available iterative DirectDensity™ (Siemens Healthineers) reconstruction algorithm (F30 kernel, strength 3) (15) the CT numbers must be converted into densities by an energy dependent conversion curve. A recently developed algorithm directly reconstructs CT projection data into relative electron densities which eliminates the use of separate conversion curves for different X-ray tube potentials. Our work evaluates this algorithm for various cancer sites and shows its applicability in a clinical workflow. Materials and methods: The Gammex phantom with tissue mimicking inserts was scanned to characterize the CT number to density conversion curves. In total, 33 patients with various cancer sites were scanned using multiple tube potentials. All CT acquisitions were reconstructed with the standard filtered back-projection (FBP). DirectDensity™ is a projection-based algorithm that reconstructs SECT scans in RED image intensities independently from the X-ray tube potential and without additional user calibrations. For all phantom acquisitions, the X-ray tube current (mAs) was manually set to have a constant CT dose index (32 cm) of approximately 30 mGy. Compared to the clinical scanning protocols, a higher CT dose index was chosen to limit noise influences in the phantom study. Instead of a 16 cm CT dose index, the 32 cm dose index was reported for this phantom study because the shoulders were also imaged according to our clinical neurological scan protocol. In the manufacturer's neurological

scan protocol (16 cm), the patient's shoulders also exceed the limited field-of-view.

The first phantom was a head-size home-made PMMA cylinder (diameter=15.0 cm, length=28.0 cm) consisting of one central hole (diameter=2.8 cm) drilled through the entire cylinder length to fit four RMI 467 tissue characterization inserts (Gammex Inc., Middleton, WI) simultaneously. Different insert combinations were inserted in the PMMA phantom so that a total of 14 tissue mimicking inserts were scanned. The second phantom was the central bulk and the inserts of the 062M phantom (CIRS Inc., Norfolk, VA). The mean CT_{HU} numbers of each insert from both phantoms were computed in a cylindrical volume-of-interest containing seven central axial CT slices and two-dimensionally eroded with four pixels in the axial CT planes to avoid the insert edges. Symbol CT_{HU} is used to denote CT numbers in Hounsfield Units (HU). The PMMA phantom was used for the calibration of four ρ_e conversion methods, and the validation of DirectDensityTM, whereas the 062M phantom was used for the validation of all five conversion methods.

ρ_e conversion methods

Single-energy CT

This thesis chapter examines two SECT methods to convert CT_{HU} numbers to ρ_e . In the first conversion method, the cylindrical PMMA phantom was used to obtain the energy dependent CT_{HU} number to ρ_e conversion curve for the 120 kVp acquisition (called "CT2RED" onwards). The CT to ρ_e conversion curve was then applied to the 062M validation phantom.

In the second conversion method, the 120 kVp acquisition of both phantoms was reconstructed by DirectDensityTM, which should be independent of the X-ray tube potential and reconstructs directly into CT_{RED} numbers [15]. CT_{RED} denotes CT numbers in units that are linearly scalable to ρ_e using a fixed linear relationship: $\rho_e = CT_{RED}/1000 + 1$ [15]. No additional ρ_e calibration procedure was required and therefore, the ρ_e values of both phantoms were used for validation of this method.

Dual-energy CT

This thesis chapter examines three DECT based CT_{HU} number to ρ_e conversion methods. In the first method, Saito's methodology [16] was used to convert the DECT images into ρ_e . This method uses a linear fit procedure between a weighted subtraction of the 80 and 140 kVp CT images to obtain scanner and protocol specific fit parameters a , b and α

$$\rho_e = a \cdot \frac{(1 + \alpha) CT_{HU}^H - \alpha CT_{HU}^L}{1000} + b \quad (\text{Eq. 1})$$

The mean 80 and 140 kVp CT_{HU} numbers of the PMMA phantom inserts (CT_{HU}^L and CT_{HU}^H) computed in cylindrical regions-of-interest in the middle of the inserts, were used in the calibration procedure. As validation, the CT_{HU}^L and CT_{HU}^H of the 062M validation phantom were combined according to [equation 1](#). Furthermore, a denoising algorithm was applied in the resulting ρ_e images, which consisted of an edge-preserving curvature flow image filter (time step=0.05, iterations=15) implemented in the Insight Segmentation and Registration Toolkit (ITK) [17]. This implementation is denoted as ‘‘Saito-15it’’.

In the second DECT based extraction method (‘‘Hünemohr-15it’’), Hünemohr’s methodology [6] was used to calculate the ρ_e image. This methodology starts from the same equation as Saito’s method [18-20], performing a weighted subtraction of the 80 and 140 kVp CT images, but bases its calibration on one single fit parameter (c_e) according to [equation 2](#). This method is implemented in *syngo.via*’s RHO/Z package (Siemens Healthineers, VB30) and is therefore added to this study. Currently, the *syngo.via* VB30 software uses a non-system specific and default fit parameter (c_e) to calculate the ρ_e images. Our Hünemohr-15it implementation, which is based on a dedicated calibration fit, derived this c_e fit parameter for a body site specific and system specific protocol (SOMATOM Confidence RT Pro) and serves as the best-case scenario possible to achieve with *syngo.via* as commercial software. The effect of denoising was investigated in [Supplementary Material A](#).

$$\rho_e = c_e \cdot \left(\frac{CT_{HU}^H}{1000} + 1 \right) + (1 - c_e) \left(\frac{CT_{HU}^L}{1000} + 1 \right) \quad (2)$$

The third and last DECT based ρ_e extraction method used a pseudo monoenergetic CT_{HU} number to ρ_e conversion curve. Here the PMI was calculated by *syngo.via*’s Monoenergetic Plus algorithm; this method will be denoted ‘‘PMI2RED’’ onwards [21]. Following Wohlfahrt *et al.* [7], the PMI was reconstructed at 79 keV. The same calibration procedure was used to obtain the CT2RED and PMI2RED conversion curves. Both conversion curves are added as [Supplementary Material B](#) and the schematic study workflow is depicted in [Supplementary Material C](#).

Brain tumor CT scans

A mass density comparison was performed for all patients and one patient dataset was used to perform proton dose calculations on the mass density datasets.

CT protocol and post-processing

To evaluate ρ_m differences between the different ρ_e conversion methods, brain CT scans of four patients were acquired with a contrast-enhanced SECT imaging protocol and a dual-spiral 80 and 140 kVp DECT imaging protocol using the SOMATOM Confidence® RT pro CT scanner. After the contrast medium injection (iodine), a latency time of 5 minutes was considered before starting the SECT scan, after that the DECT scan was directly acquired (<1 minute). The impact of the contrast medium on the images was investigated for one of the patients (Supplementary Material D). All scans were acquired with a constant CT dose index (32 cm) of 20 mGy and reconstructed with a voxel dimension of 0.64x0.64x1.00 mm³. The study was approved by the internal review board of Maastric Clinic (P0243).

The SECT and DECT acquisitions were reconstructed with the SAFIRE technique with beam hardening correction for bone enabled (Q34 kernel, strength 3). Additionally, the SECT-based DirectDensity™ reconstruction (F30 kernel, strength 3) was performed for one representative patient, since this algorithm is not recommended for proton therapy³. The DECT reconstructions were registered using deformable image registration software (Elastix) [22], and post-processed to calculate the ρ_e images based on Saito's and Hünemohr's methods. Additionally, *syngo.via*'s DECT software package was used to calculate the PMIs at 79 keV.

A proton treatment plan was created for one of the four brain CT scan with a fictitious tumor and prior to the planning stage, the five ρ_e images were converted to ρ_m by using a linear fit relationship (maximum fit error <1%) based on human tissues and recommended in TG186 [9]: $\rho_m = -0.175 + 1.176 \cdot \rho_e$ ($R^2 = 0.99992$). Between the CT acquisitions, the anatomy of the patient can change due to motion, therefore all ρ_m images originating from the DECT conversion methods were elastically deformed to the ρ_m image originating from the SECT 120 kVp scan. Finally, the same values of ρ_m for the voxels outside the body structure were assigned in all sets to assure that the differences found in the treatment plan were caused only by the ρ_m values of the patient geometry.

Proton treatment planning

The Clinical Target Volume (CTV) corresponding to a fictitious brain tumor (cranio-pharyngioma), as well as the OARs were delineated by a radiation oncologist in the conventional 120 kVp SECT scan. A pencil beam scanning treatment plan was made in RayStation v8A Monte Carlo (RaySearch Laboratories AB, Stockholm, Sweden) using the CT2RED ρ_m -image. Three coplanar fields (one posterior and two lateral) were aimed at the CTV with a prescription dose of 54 Gy (RBE) in 30 fractions (**Figure 1**). The beam angles (80°, 180°, 290°) were chosen to avoid crossing the right and

3 private correspondence with Siemens Healthineers

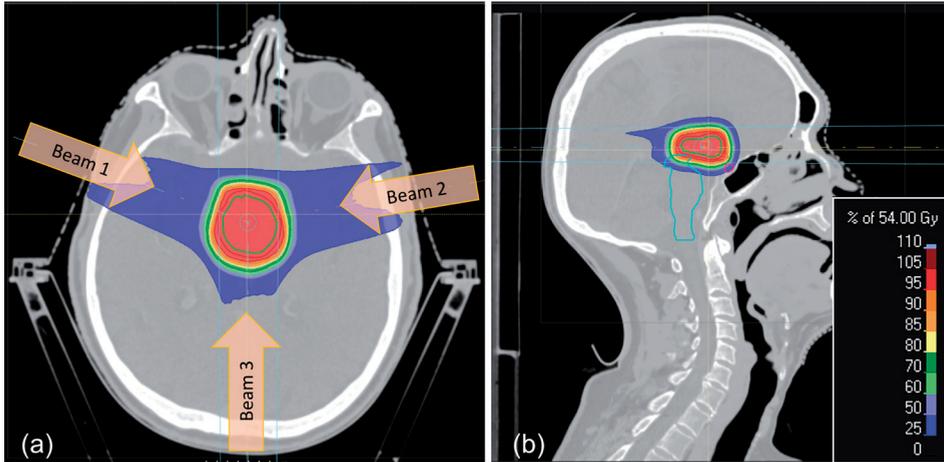


Figure 1. A proton treatment plan with a fictitious brain tumor irradiated by three proton beams was created to evaluate dose and proton range differences for one example patient.

left hippocampi, which have a low dose tolerance ($D_{40\%} \leq 7.3$ Gy) [23]. The optical chiasm with a tolerance of $D_{0.03cc} \leq 55$ Gy is located at the distal edge of the posterior beam [23].

The optimization toolbox and the RayStation Monte Carlo dose engine were used to compute the final dose with 0.5% statistical uncertainty in the CT2RED ρ_m image. After this, the final plan was recalculated, without performing a new optimization, in the other four sets of density images (DirectDensityTM, PMI2RED, Hünemohr-15it and Saito-15it) and relative differences between the CT2RED and the other dose distributions were assessed. Furthermore, dose volume histograms (DVH) for the CTV and selected OARs (optical chiasm, brainstem, hippocampi and pituitary gland) were analyzed. Finally, a plan with a single posterior beam (beam 3 in Figure 1) was made with the purpose of quantifying differences in proton range (defined as the 80% distal fall-off) between the different methods. Beam 3 was chosen to investigate the dose influence on the chiasm.

Results

Calculated relative electron density accuracy in phantoms

The relative difference between the determined ρ_e by the five conversion methods and the certified ρ_e for the PMMA and 062M phantoms is illustrated in Figure 2(a). In both phantoms, the largest ρ_e errors were found in the high-density bone inserts for the DirectDensityTM image reconstruction method. Compared to the manufacturer's ρ_e certificate, DirectDensityTM overestimated the ρ_e of 'Bone 1.53' by 3.9% and 'Bone

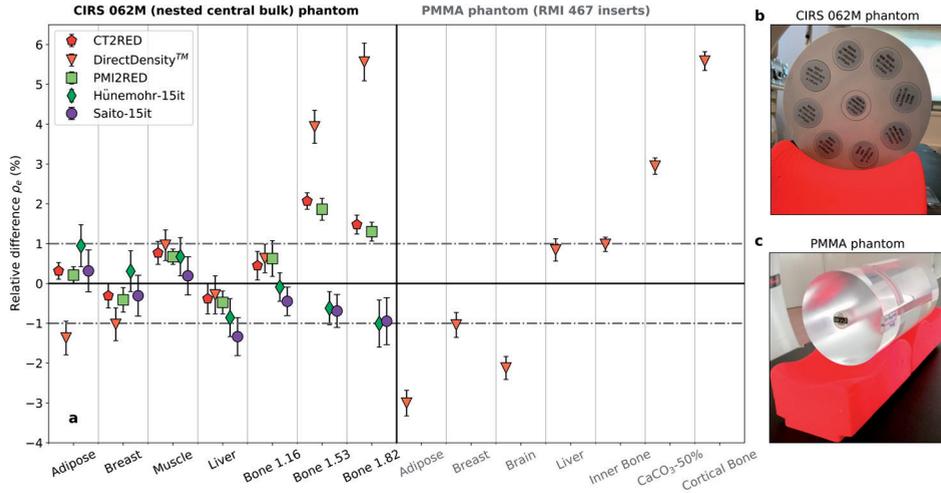


Figure 2. The relative ρ_e difference (%) (measurement-reference) for the five conversion methods. The error bars indicate ± 1 standard deviation (a) On the left, the differences were calculated for the 062M phantom. On the right, the differences were calculated for the PMMA phantom. The measurement setup is illustrated in (b) and (c).

1.82' by 5.5%. The lowest root-mean-square (RMS) errors for ρ_e were found for the Saito-15it (a=0.999, b=0.995, $\alpha=0.864$) and the Hünemohr-15it ($c_e=0.873$) methods, which were below $\pm 1.0\%$. Overall, the RMS errors for the 062M inserts were 2.7% for DirectDensity™ and 1.0% for CT2RED and PMI2RED.

The standard deviations calculated in the inserts for the Saito-15it and the Hunemohr-15it methods were larger compared to the other conversion methods, although edge-preserving denoising was applied as post-processing (Supplementary Material A). Between these two DECT methods, no significant differences in RED calculation were observed. ($p=0.003$, paired t-test). The CT2RED and PMI2RED methods showed comparable errors; only in the medium and high-density bone inserts, the ρ_e errors were larger than 1.4% for both methods.

Mass density differences in the patient geometries

Absolute ρ_m differences (g/cm^3) in arbitrary cranial slices between the conventional CT2RED conversion method and three other methods are depicted in Figure 3 for the different brain CT scans. In the right column of Figure 3, a histogram indicates the normalized frequency distribution of the absolute ρ_m differences in bins of $0.01 \text{ g}/\text{cm}^3$ between the conventional CT2RED method (as reference) and the PMI2RED, Hünemohr-15it and Saito-15it conversion methods. For each of the four brain CT scans, Figure 4 shows the average ρ_m differences for all tissues (excl. air cavities and dental filling artifacts), but also for the soft tissues ($0.9 < \rho_m < 1.2 \text{ g}/\text{cm}^3$) and for

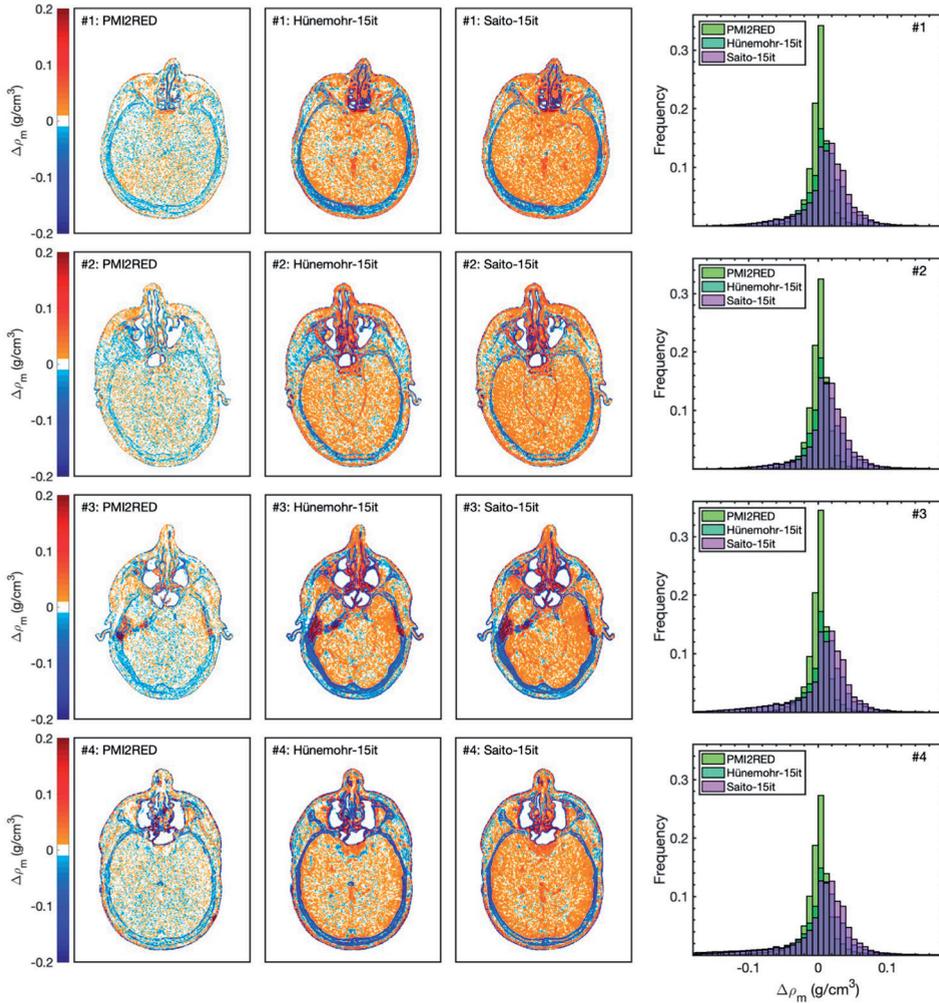


Figure 3. Absolute mass density differences ($\Delta\rho_m$) are shown in one representative CT slice for each of the four patients (#1-4). Differences are calculated for the PMI2RED, Hünemohr-15it and Saito-15it conversion methods minus the conventional reference CT2RED. In the right panels, the frequency histogram is shown over the entire CT image excluding air cavities and dental artifacts.

bone tissues ($\rho_m \geq 1.2 \text{ g/cm}^3$) separately. Air cavities were excluded by thresholding ($\rho_m < 0.2 \text{ g/cm}^3$) and dental artifacts were excluded by manual artifact segmentation. The DirectDensityTM results for patient #1 are shown in [Supplementary Material E](#).

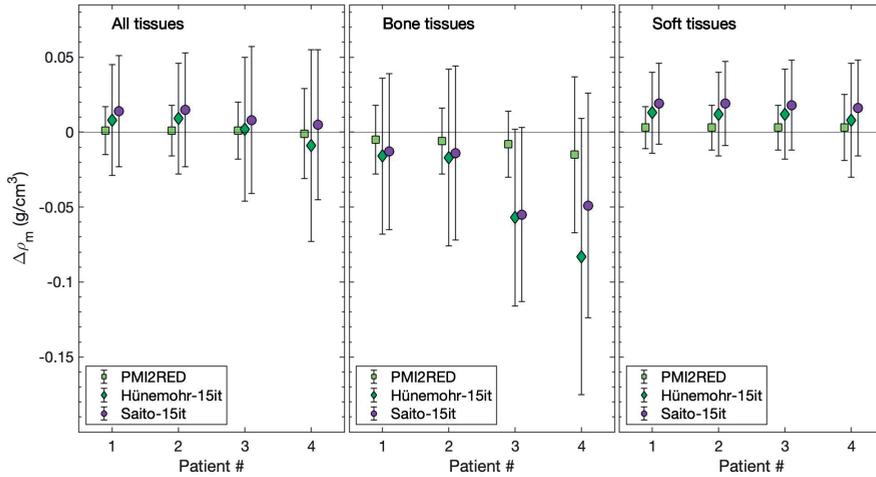


Figure 4. Mass density differences ($\Delta\rho_m$) (± 1 standard deviation) for the PMI2RED, Hünemohr-15it and Saito-15it methods minus the conventional CT2RED conversion method calculated over the entire patient geometry excluding air and dental fillings, for all tissues (left), only for bone tissues (middle) and only for soft tissues (right panel).

With respect to CT2RED, the PMI2RED (-0.01 ± 0.01 g/cm³), Hünemohr-15it (-0.04 ± 0.04 g/cm³) and Saito-15it (-0.02 ± 0.02 g/cm³) conversion methods resulted in overall lower bone densities (i.e. negative differences compared to CT2RED) averaged over the four brain CT scans. For the same conversion methods, overall positive differences were observed for the soft tissues: 0.003 ± 0.001 g/cm³, 0.011 ± 0.002 g/cm³ and 0.018 ± 0.001 g/cm³, respectively.

Differences in proton dose distributions

Relative dose differences between the CT2RED and the other three conversion methods for a representative axial CT image slice of the transversal plane are shown in [Figure 5\(a-c\)](#). The plan evaluated on the Saito-15it image shows the largest differences compared to the initial plan (down to -6%), followed by the Hunemohr-15it method (down to -4%) and the PMI2RED method (almost no visible differences). The DVHs resulting from each plan are depicted in [Figure 5\(d\)](#) for the CTV and five selected OARs, with the DVH of the CTV zoomed-in on [Figure 5\(e\)](#). For the dose from one posterior beam, differences in proton range resulted in an average of -0.1 ± 0.3 mm for PMI2RED; -0.8 ± 0.4 mm for Hünemohr-15it, and -1.2 ± 0.4 mm for Saito-15it.

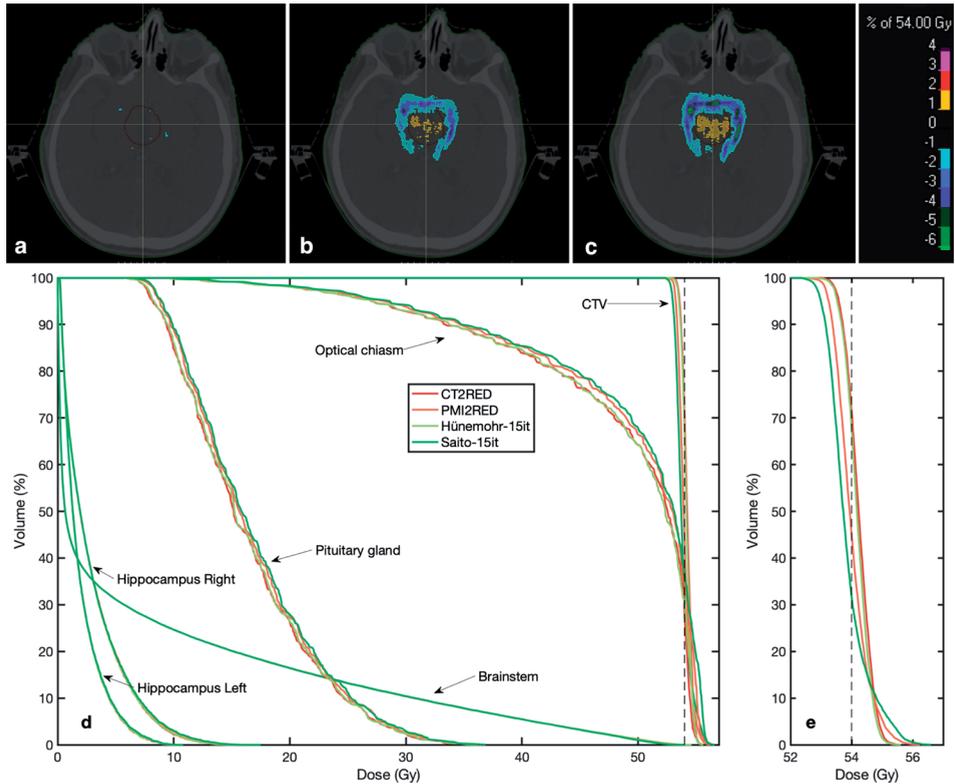


Figure 5. Top panel: dose differences between the plan calculated on the CT2RED images and the same plan evaluated on (a) PMI2RED, (b) Hünemohr-15it and (c) Saito-15it datasets. Bottom panel: (d) dose volume histograms of the five evaluated methods for the CTV (solid line) and five organs-at-risk; (e) zoomed-in CTV dose volume histogram. The method specific DVH curves for the brainstem largely overlap.

Discussion

The recent availability of CT scanners with DECT capabilities in addition to the standard SECT mode, and different methods of converting CT_{HU} or CT_{RED} numbers to ρ_e in radiotherapy facilities raises the question which one to use and its impact on treatment planning. This study aims to investigate the differences in relative electron densities and dose calculations for proton treatment planning depending on which conversion method is used: CT2RED, PMI2RED, Hünemohr-15it and Saito-15it, from which the first two methods are part of commercially available software.

Initially, the ρ_e accuracy was assessed in two phantoms. The 062M phantom was used as validation phantom for all methods and relative errors with respect to the

reference ρ_e showed a higher discrepancy in the high-density materials (Figure 1). Saito-15it and Hunemohr-15it had the lowest root-mean-square (RMS) error of 0.7%, whereas DirectDensity™ showed the highest RMS error of 2.7%. This study used head-sized phantoms, but it should be noted that the DirectDensity™ algorithm does not support user calibration and is developed by Siemens Healthineers on an abdominal-sized phantom⁴. It is expected that the RED accuracy of the DirectDensity™ will improve when the algorithm is optimized for head-sized volumes, however this remains speculation and needs further investigation.

With SECT and DECT scans from four patients, ρ_e images were created using each conversion method and converted in ρ_m using the recommendation from TG186 [9]. Lacking a ground truth for the patient scans, the CT2RED ρ_m images were used as reference because this method is used at our institute for photon (Eclipse TPS, Varian Medical Systems) and for proton dose calculations. Differences with the other conversion methods were assessed (Figure 3 and Figure 4).

The Hünemohr-15it method aims to reproduce the results that the commercial *syngo.via* RHO/Z package would give, if a proper scanner-, size- and energy-specific calibration were used. Scanner and body site specific calibration parameters could not be adapted in the *syngo.via* version installed at our institution, therefore our own implementation of the method had to be used. The use of this commercial RHO/Z package without an optimal calibration might result in relative electron density discrepancies and dose distribution differences. The ability to specify scan-specific fit parameters in clinical software would be beneficial for the accuracy of the relative electron density images and furthermore the radiotherapy dose calculations.

The dose distribution discrepancies caused by the ρ_m differences were investigated for one example patient, in which the CT2RED derived dataset was used as reference. In the planning stage of this treatment plan, robust planning [24] was considered to follow the trend in clinical practice to account for set-up and range uncertainties, but it was found that a robust plan did not alter the main conclusions shown. Therefore, no robust optimization was used in this study. Results in Figures 5(a-c) agree with the ρ_m differences observed in Figures 3 and Figure 4. The contributions of bone and soft tissues to these dose differences were assessed and it was found that, although dense bone tissues show the higher absolute ρ_m differences between methods, it has a negligible contribution for this patient plan because the proton beam path through brain tissue is much longer than the beam path through bone.

For the single posterior beam plan (beam 3), average 80% distal fall-off range differences were reported between -1.2 mm (Saito-15it) and 0.1 mm (PMI2RED) which fall within the voxel size of (1 mm)³ typically used in radiotherapy for head CT

⁴ Private correspondence with Siemens Healthineers.

scans. The 0.1 ± 0.3 mm difference found with the PMI2RED conversion is lower but comparable to the 0.3 mm range reported for a head-and-neck cancer patient by Wohlfahrt *et al.* [7] where the hypothetical treatment beam had to pass more bone than in our treatment plan.

The optical chiasm and pituitary gland show the largest differences in DVH for the three methods, because of their location in the distal edge of the posterior beam and their small volume, making them more sensitive to small variations. In contrast, the brainstem and hippocampi show similar dose coverages for all methods, since these organs are not in the beam paths. To follow up on this study, a larger cohort of patients and different sites should be analysed to confirm the robustness and consistency of each method.

Given the wide choice of commercially available methods for converting SECT and DECT images to ρ_e , studies are needed to assess the accuracy of these methods, as well as its impact for proton dose calculations. However, there is still a gap between performing accuracy studies in certified ρ_e phantoms with known compositions and applying these methods in human CT images. Furthermore, full-DECT treatment planning is currently not an option in commercial TPS, where the only solution for the user is to pre-process the DECT images and use the resulting ρ_m (or SPR) image as input [25]. This work quantifies the differences in ρ_m for four patients and quantifies differences in dose and proton range for a neurological patient between the conventional CT2RED method and three other SECT- or DECT- based conversion methods. For this treatment site, the PMI2RED method was equivalent to the conventional CT2RED method in terms of dose distribution, CTV coverage and dose to the OARs, whereas Hünemohr-15it and Saito-15it led to differences in proton dose and range. Which method approximates best reality in a patient cannot be determined in this study. However, the mass density differences observed in our patient cohort were in agreement with the differences observed in the phantom study. In this phantom study, DECT-based methods with body site and protocol specific calibration such as Hünemohr-15it and Saito-15it achieved the highest relative electron density accuracy, which allows to speculate that these methods also achieved the highest accuracy in our patient cohort. Furthermore, the dose differences between the different conversions were in agreement with previously published research [4, 26].

Not only further research is required to evaluate the accuracy of the available conversion methods in larger patient cohorts or in realistic phantom geometries with known material properties, also efforts from the imaging and TPS manufacturers are required to improve the accuracy of the radiotherapy dose calculations. Clinics should be enabled to set optimized DECT parameters (e.g. c_e fit coefficient in Hünemohr-15it) according to their clinical scan protocols and should be enabled to directly use post-processed DECT acquisitions in the TPS to perform dose calculations.

References

1. Jakobi, A., et al., *Identification of Patient Benefit From Proton Therapy for Advanced Head and Neck Cancer Patients Based on Individual and Subgroup Normal Tissue Complication Probability Analysis*. *Int J Radiat Oncol Biol Phys*, 2015. **92**(5): p. 1165-1174.
2. van de Water, T.A., et al., *The potential benefit of radiotherapy with protons in head and neck cancer with respect to normal tissue sparing: a systematic review of literature*. *Oncologist*, 2011. **16**(3): p. 366-77.
3. Weber, D.C., et al., *RapidArc, intensity modulated photon and proton techniques for recurrent prostate cancer in previously irradiated patients: a treatment planning comparison study*. *Radiat Oncol*, 2009. **4**: p. 34.
4. Almeida, I.P., et al., *Monte Carlo proton dose calculations using a radiotherapy specific dual-energy CT scanner for tissue segmentation and range assessment*. *Phys Med Biol*, 2018. **63**(11): p. 115008.
5. Hudobivnik, N., et al., *Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images*. *Med Phys*, 2016. **43**(1): p. 495.
6. Hunemohr, N., et al., *Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates*. *Phys Med Biol*, 2014. **59**(1): p. 83-96.
7. Wohlfahrt, P., et al., *Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans*. *Int J Radiat Oncol Biol Phys*, 2017. **97**(2): p. 427-434.
8. Wohlfahrt, P., et al., *Dual-energy CT based proton range prediction in head and pelvic tumor patients*. *Radiation Oncol*, 2017. **125**(3): p. 526-533.
9. Beaulieu, L., et al., *Report of the Task Group 186 on model-based dose calculation methods in brachytherapy beyond the TG-43 formalism: current status and recommendations for clinical implementation*. *Med Phys*, 2012. **39**(10): p. 6208-36.
10. Taasti, V.T., et al., *Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples*. *Phys Med Biol*, 2017. **63**(1): p. 015012.
11. Constantinou, C., J.C. Harrington, and L.A. DeWerd, *An electron density calibration phantom for CT-based treatment planning computers*. *Med Phys*, 1992. **19**(2): p. 325-7.
12. Das, I.J., et al., *Computed tomography imaging parameters for inhomogeneity correction in radiation treatment planning*. *J Med Phys*, 2016. **41**(1): p. 3-11.
13. Schneider, U., E. Pedroni, and A. Lomax, *The calibration of CT Hounsfield units for radiotherapy treatment planning*. *Phys Med Biol*, 1996. **41**(1): p. 111-24.
14. Goma, C., I.P. Almeida, and F. Verhaegen, *Revisiting the single-energy CT calibration for proton therapy treatment planning: a critical look at the stoichiometric method*. *Phys Med Biol*, 2018. **63**(23): p. 235011.
15. van der Heyden, B., et al., *Clinical evaluation of a novel CT image reconstruction algorithm for direct dose calculations*. *Physics and Imaging in Radiation Oncology*, 2017. **2**: p. 11-16.
16. Saito, M., *Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship*. *Med Phys*, 2012. **39**(4): p. 2021-30.
17. Malladi, R. and J.A. Sethian, *Image processing: flows under min/max curvature and mean curvature*. *Graphics Models and Image Processing*, 1996. **58**(2): p. 127-141.
18. Almeida, I.P., et al., *Reply to: "Comment on: Dual-energy CT quantitative imag-*

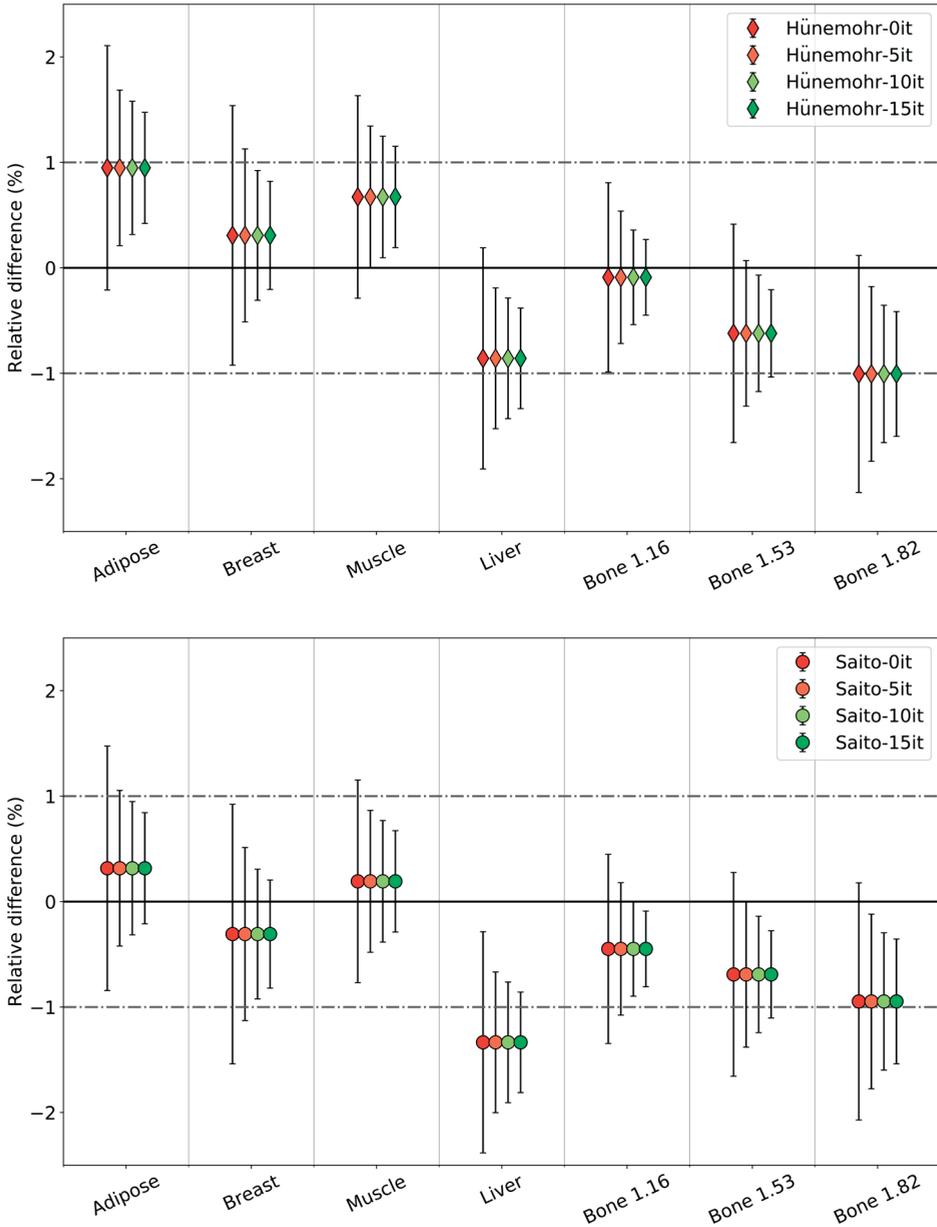
- ing: A comparison study between twin-beam and dual-source CT scanners [Med. Phys. 44(1), 171-179 (2017)]". Medical Physics, 2018. **45**(8): p. 3997-3998.
19. Almeida, I.P., et al., *Dual-energy CT quantitative imaging: a comparison study between twin-beam and dual-source CT scanners*. Med Phys, 2017. **44**(1): p. 171-179.
 20. Saito, M., *Comment on: Dual-energy CT quantitative imaging: A comparison study between twin-beam and dual-source CT scanners [Med. Phys. 44(1), 171-179 (2017)]*. Medical Physics, 2018. **45**(8): p. 3996-3996.
 21. Grant, K.L., et al., *Assessment of an advanced image-based technique to calculate virtual monoenergetic computed tomographic images from a dual-energy examination to improve contrast-to-noise ratio in examinations using iodinated contrast media*. Invest Radiol, 2014. **49**(9): p. 586-92.
 22. Klein, S., et al., *elastix: a toolbox for intensity-based medical image registration*. IEEE Trans Med Imaging, 2010. **29**(1): p. 196-205.
 23. Lambrecht, M., et al., *Radiation dose constraints for organs at risk in neuro-oncology; the European Particle Therapy Network consensus*. Radiother Oncol, 2018. **128**(1): p. 26-36.
 24. Unkelbach, J. and H. Paganetti, *Robust Proton Treatment Planning: Physical and Biological Optimization*. Semin Radiat Oncol, 2018. **28**(2): p. 88-96.
 25. Wohlfahrt, P., et al., *Evaluation of Stopping-Power Prediction by Dual- and Single-Energy Computed Tomography in an Anthropomorphic Ground-Truth Phantom*. Int J Radiat Oncol Biol Phys, 2018. **100**(1): p. 244-253.
 26. Bar, E., et al., *Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples*. Med Phys, 2018. **45**(1): p. 48-59.

Supplementary Material

A. Impact of noise in Saito's and Hünemohr's method

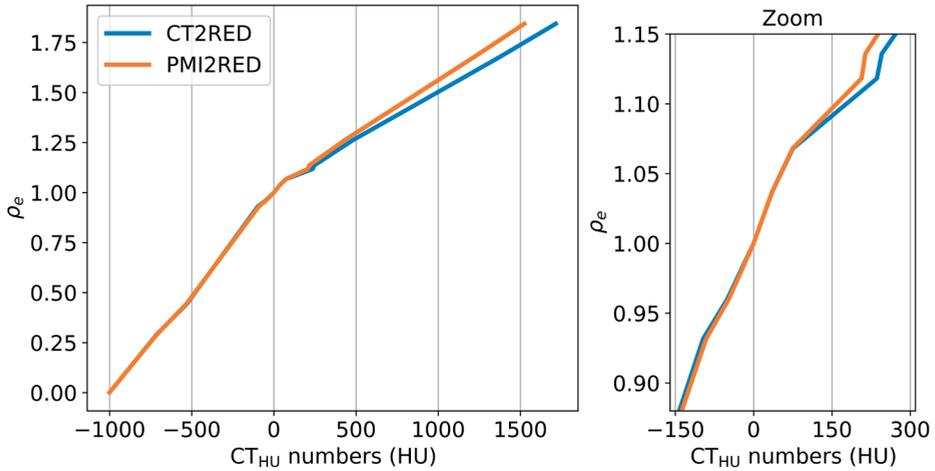
Saito's and Hünemohr's methodologies were used to convert the DECT images into ρ_e . Both methods combine the 80 and 140 kVp CT datasets according to [Equation 1](#) and [Equation 2](#) in this thesis chapter. To reduce noise influences after combining the 80 kVp and the 140 kVp CT datasets, an edge preserving denoising algorithm was applied to the calculated ρ_e images, which consisted of an edge-preserving curvature flow image filter implemented in the Insight Segmentation and Registration Toolkit (ITK). The influence of the iteration number on the ρ_e accuracy was investigated for 3 different iteration numbers: 5, 10, and 15 iterations.

After applying the edge-preserving denoising algorithm as a post-processing step, the relative differences were computed between the ρ_e from the reference and the ones obtained by the two methods. For the 062M tissue mimicking inserts, the mean value remained the same after the post-processing step, while the standard deviations reduced by 50% between Saito-0it and Saito-15it or Hünemohr-0it and Hünemohr-15it.



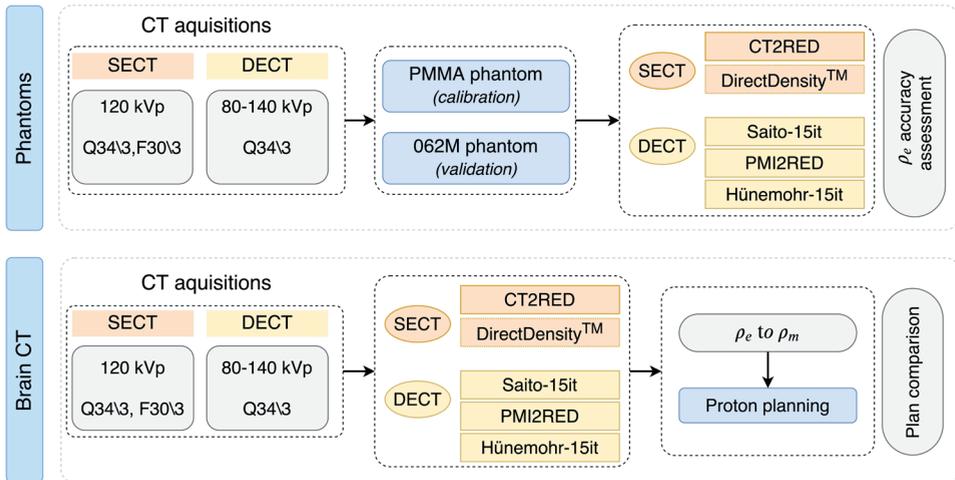
Supplementary 1. Relative difference (%) \pm 1 standard deviation (measured-reference) in ρ_c for two methodologies and for different iteration values of the edge-preserving curvature flow image filter in the post-processing step (0, 5, 10 and 15). Zero iterations correspond to no denoising.

B. CT2ED and PMI2ED conversion curves



Supplementary 2. Relative electron density conversion curves used to convert 120 kVp CT and 79 keV PMI images in Hounsfield Units to relative electron densities (RED).

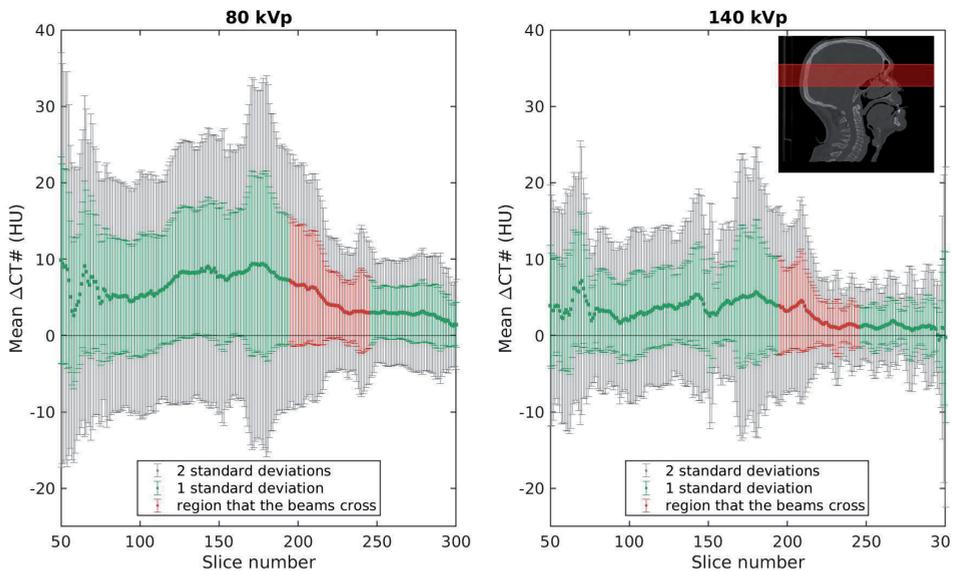
C. Study workflow



Supplementary 3. Workflow of the phantom and the patient brain CT acquisitions. The phantom study aims to assess relative electron density differences between five CT to density conversion methods. The patient study investigates density and proton dose differences between all conversion methods in one patient treatment plan.

D. Impact of the presence of contrast medium in the patient scans

In clinical practice, when a contrast-enhanced CT is desired for tumor and organ-at-risk delineation, a second scan without contrast is acquired for proton treatment planning. However, this study was part of a clinical trial where multiple CT scans were acquired from one patient; a DECT scan prior to contrast injection, a SECT scan after contrast injection and a DECT scan after contrast injection. For the proton dose calculations presented in the chapter, only the contrast enhanced SECT and contrast-enhanced DECT scans were used. The individual non-contrast DECT images of 80 kVp and 140 kVp were only used for the comparison with the contrast-enhanced 80 kVp and 140 kVp CT images. The mean Hounsfield Unit (HU) differences between the contrast-enhanced images and the non-contrast images are shown in [Suppl. 4](#), highlighting the slices through which the beams crossed. In these slices of interest, the maximum mean CT number difference was 7 HU for the 80 kVp CT scans and 5 HU for the 140 kVp CT scans.



Supplementary 4. Mean difference in CT numbers (1 and 2 standard deviations) between the contrast-enhanced and the non-contrast CT images, for 80 kVp (left) and 140 kVp (right). The top right panel: sagittal CT plane of the patient illustrating with a red box the slices crossed by the three proton beams shown in Figure 1. The same slices are highlighted in both plots.

An additional phantom study, supported by patient data, was performed to evaluate the performance of the CT conversion methods on iodine inserts. The dose differences reported in this study were evaluated on a patient treatment plan with a virtual non-contrast-enhanced tumor. However, the patient had a real brain tumor close to the cranium of which the mean 80kVp HUs of brain tissue and the tumor were determined as extreme scenario. The 80 kVp CT reconstruction was picked because the photo-electric absorption of X-rays is sensitive to changes in material composition, especially in high-Z materials (iodine) and at lower energies (80 kVp). Then, the iodine insert with the highest similarity in measured HUs (84 ± 7 HU), the 2 mg/ml Gammex insert, was chosen for further analysis. The atomic weight fractions of the 2 mg/ml iodine insert are listed in Table S1.

Table S1.

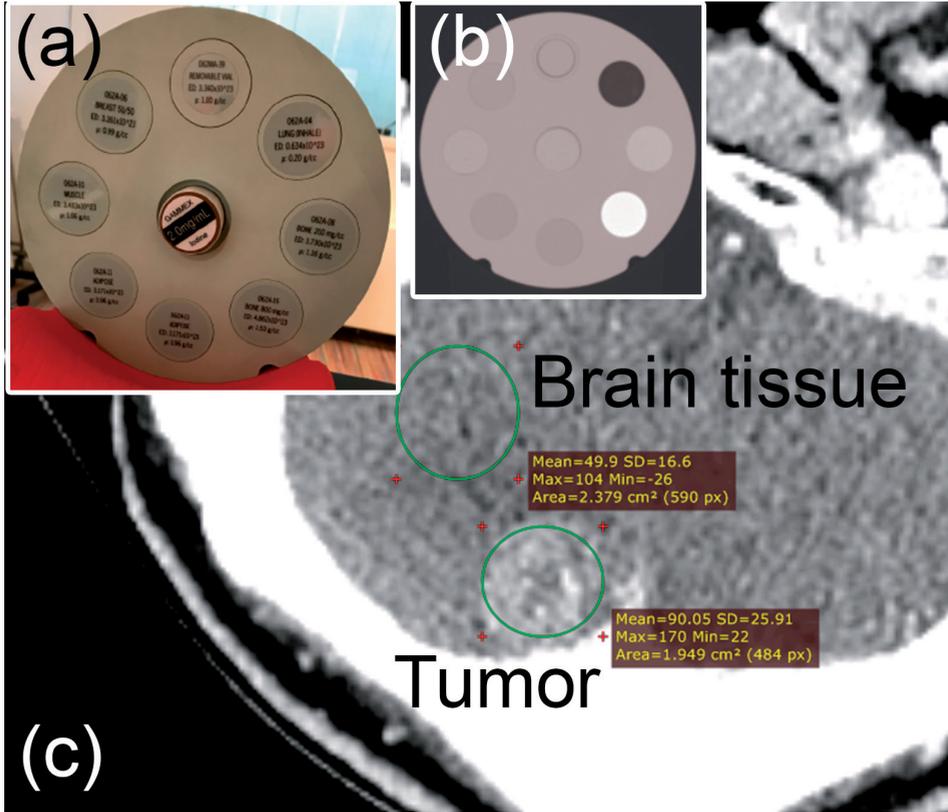
	ρ_m	ρ_e	H	O	C	N	Cl	Ca	Mg	I
Iodine insert	1.03	1.00	8.64	17.51	69.53	2.15	0.13	1.81	0.03	0.20

The iodine insert was placed in the middle of the central bulk of the 062M phantom (Suppl. 5(a) and 5(b)). 120 kVp SECT imaging was performed to evaluate the accuracy of the CT2RED conversion method and the DirectDensity algorithm. 80/140 kVp dual-spiral DECT acquisitions were acquired and processed to evaluate the PMI2RED, the Saito-15it, and the Hünemohr-15it conversion methods. The accuracies of the different relative electron density conversion methods are listed in Table S2 (measurement – reference)/reference.

Table S2.

Conversion method	(Measurement-Ref)/Ref * 100%
CT2RED	4.8%
DirectDensity TM	4.1%
PMI2RED	4.3%
Hünemohr-15it	0.2%
Saito-15it	1.0%

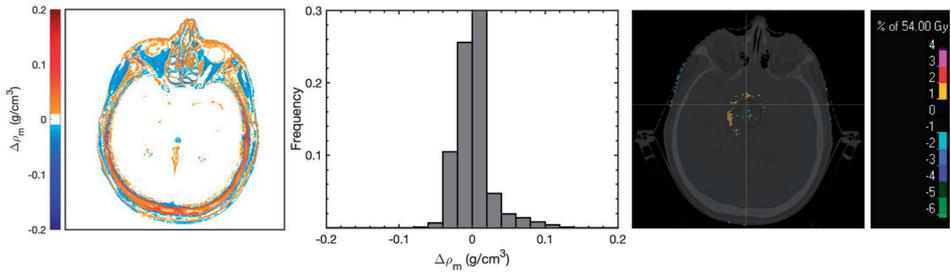
The HU difference between the 2.0 mg/ml iodine insert and the phantom bulk was approximately 80HU, as extreme example. Differences found in the patient example between contrast-enhanced and non-contrast 80 kVp images showed HU differences of maximum 7HU. Converted to RED, this only introduces a small RED error.



Supplementary 5. (a) imaging setup of the 062M phantom (CIRS) with the iodine insert 2.0 mg/ml in the middle, (b) depicts the SECT scan of the imaging setup acquired at 80 kVp, (c) shows the real patient tumor (not the virtual tumor used for dose calculations in this study) in a 80 kVp CT reconstruction where two regions-of-interest were drawn to investigate the mean absolute Hounsfield Units for brain tissue (50±17 HU) and the tumor (90±26 HU).

E. DirectDensity™ results for patient #1

The CT acquisitions for patient #1 were reconstructed using DirectDensity™. The resulting ρ_e -images were converted to ρ_m as explained in subsection “CT protocol and post-processing”. The average ρ_m difference (± 1 standard deviation) over the whole CT scan between the DirectDensity™ image and the CT2RED image was $0.00 \pm 0.02 \text{ g/cm}^3$. The dose differences for the plan calculated in the CT2RED set and recalculated with DirectDensity™ were within 1%. **Suppl. 6** summarizes the results for DirectDensity™: absolute ρ_m differences with respect to the CT2RED image for a representative slice, the frequency distribution calculated over all slices (excl. air cavities and dental filling artifacts) and the resulting dose difference for the proton treatment plan with three beams.



Supplementary 6. Left panel: Absolute mass-density differences ($\Delta\rho_m$) between DirectDensity™ and the conventional CT2RED conversion method. Middle panel: Probability histogram calculated over the entire CT (excl. air cavities and dental filling artifacts). Right panel: dose differences between the plan calculated on the CT2RED images and the same plan evaluated on the DirectDensity™ set.

5

Chapter 5

Dual-energy CT for automatic organs-at-risk segmentation in brain-tumor patients using a multi-atlas and deep learning approach

Brent van der Heyden*, Patrick Wohlfahrt*, Danielle BP Eekers, Christian Richter, Karin Terhaag, Esther GC Troost, and Frank Verhaegen

* Equal contribution

Published in

Scientific Reports (2019)

<https://doi.org/10.1038/s41598-019-40584-9>

Abstract

In radiotherapy, computed tomography (CT) datasets are mostly used for radiation treatment planning to achieve a high-conformal tumor coverage while optimally sparing healthy tissue surrounding the tumor, referred to as organs-at-risk (OARs). Based on CT scan and/or magnetic resonance images, OARs have to be manually delineated by clinicians, which is one of the most time-consuming tasks in the clinical workflow. Recent multi-atlas (MA) or deep learning (DL) based methods aim to improve the clinical routine by an automatic segmentation of OARs on a CT dataset. However, so far, no studies investigated the performance of these MA or DL methods on dual-energy CT (DECT) datasets, which have been shown to improve the image quality compared to conventional 120 kVp single-energy CT. In this study, the performance of an in-house developed MA and a DL method (two-step three-dimensional U-net) was quantitatively and qualitatively evaluated on various DECT-derived pseudo-monoenergetic CT datasets ranging from 40 keV to 170 keV. At lower energies, the MA method resulted in more accurate OAR segmentations. Both the qualitative and quantitative metric analysis showed that the DL approach often performed better than the MA method.

Conflict of Interest

This work was partially performed with research funding from Varian Medical Systems (Palo Alto, USA), Siemens Healthineers (Forchheim, Germany), and the NVIDIA Corporation (Santa Clara, USA).

Acknowledgements

The authors would like to thank Dr. Xue Feng from the University of Virginia for providing valuable information about the deep learning algorithm used in this work, and for making the algorithm publicly available to the community: https://github.com/xf4j/aapm_thoracic_challenge.

Introduction

In the clinical radiotherapy workflow, the targeted tumor volume and surrounding organs-at-risk (OARs) are manually delineated on image datasets derived from computed tomography (CT), often in combination with magnetic resonance imaging (MRI). In current clinical practice, CT acquisitions are mandatory to calculate radiation treatment plans and to interpret dose evaluation metrics [1, 2]. As manual delineation is one of the most time-consuming tasks and subject to inter- and intra-observer variability, a considerable interest in automatic delineation has been seen in recent years to further improve this well-recognized source of uncertainty in the radiation planning process [3-6]. In the last years, multi-atlas (MA) or deep learning (DL) methods have been investigated for automatic image segmentation [3]. Both approaches use a set of labeled medical image datasets as input for model training and finally application. In this study, CT image datasets with manually delineated OARs serve as reference. Although the fact that automatic contouring algorithms are commercially available, their use in radiotherapy clinics remains limited.

The performance of MA and DL methods for automatic contouring has been already investigated on CT or MRI datasets [3, 7, 8]. However, to our knowledge, so far no studies explored the use of such methods on dual-energy CT (DECT) image datasets, which provide additional tissue information contributing to a reduction of the intra-observer variability of physicians [9]. DECT scans consist of two CT datasets acquired with different X-ray spectra or energy separation on the detector level. The combined single DECT datasets can be used to calculate a pseudo-monoenergetic image (PMI) with a weighted sum of the low- and high-energy CT scan [10, 11]. To enhance the image quality, some commercial systems additionally include noise-suppression algorithms [12]. Furthermore, it has been demonstrated that a PMI can have superior image quality compared to 120 kVp single-energy CT (SECT) [12-15], which is the current clinical standard in most radiotherapy facilities. The influence of beam hardening on CT numbers can be reduced by PMI datasets leading to improvements in radiation treatment planning, especially in proton therapy [16]. A PMI can also contribute to suppress metal artifacts in CT imaging [10, 17].

This study first aims to quantitatively evaluate pseudo-monoenergetic CT datasets of different photon energies ranging from 40 keV to 170 keV for OAR segmentation in primary brain-tumor patients using an in-house developed 3D MA and 3D DL based image segmentation method. Secondly, two experienced radiation oncologists and one experienced radiation technologist performed a qualitative scoring to assess the clinical relevance and accuracy of automatic OAR segmentation. For this evaluation, both methods were applied on two pseudo-monoenergetic CT datasets of different energy (40 keV and 70 keV).

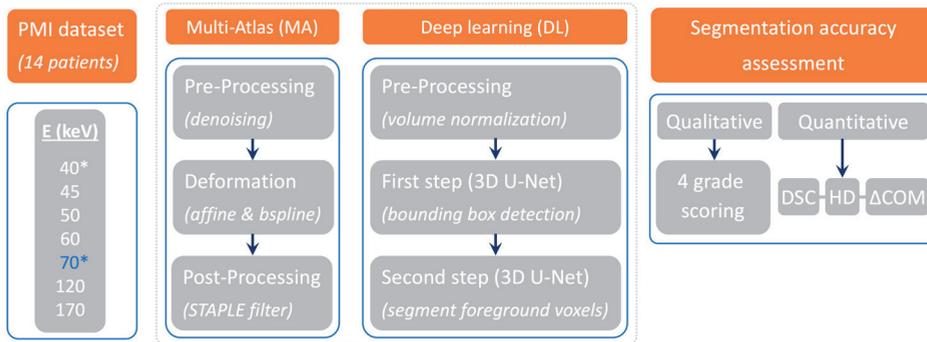


Figure 1. Study flowchart. The multi-atlas (MA) method was applied on all pseudo-monoenergetic image (PMI) datasets. The deep learning (DL) approach was applied on the energies indicated with an asterisk. The reference energy E (70 keV) was used for manual contouring of the organs-at-risk (OARs). The quantitative and qualitative segmentation accuracy was assessed between the automatically generated contours and the manual contour using the dice similarity coefficient (DSC), the 95th percentile Hausdorff Distance (HD), the center of mass displacement ΔCOM and a four-grade scoring system.

Materials and Methods

Patient cohort and DECT imaging

For this retrospective study following the regulatory guidelines and approved by the local ethics committee (EK535122015, University Hospital Carl Gustav Carus, Dresden), 14 primary brain-tumor patients were randomly selected. Each patient agreed with an informed consent to use their pseudonymized and anonymized data for scientific purposes (Figure 1). All patients underwent DECT imaging (80/140 kVp) for proton treatment planning at University Proton Therapy Dresden (Dresden, Germany) with a single-source DECT scanner (SOMATOM Definition AS, Siemens Healthineers, Forchheim, Germany) [18]. Each DECT scan was acquired with a constant CT dose index (32 cm) of 20.8 mGy and reconstructed using an iterative reconstruction algorithm including a beam hardening correction for bone (SAFIRE, Q34/3 kernel, both Siemens terminology) with a voxel size of $0.98 \times 0.98 \times 2.00 \text{ mm}^3$. PMI datasets of 7 different energies (40, 45, 50, 60, 70, 120 and 170 keV) were generated from the 80 kVp and 140 kVp DECT scans using the application SYNGO.CT DE MONOENERGETIC PLUS of the image post-processing software SYNGO.VIA (Siemens Healthineers, Forchheim, Germany). The input PMI dataset of the segmentation methods will be referred to as, e.g. PMI-40, i.e. the PMI dataset calculated at an energy of 40 keV. An experienced radiation oncologist used the PMI-70 dataset to delineate OARs for radiation treatments of head-and-neck and neuro-oncology cases, such as the brainstem, eyes, lenses, optic nerves and parotid glands.

Automatic image segmentation methods

Two fundamentally different image segmentation methods are investigated; (i) an organ-driven MA method and (ii) a two-step 3D U-Net DL method. In 2017, both segmentation methods were evaluated with state-of-the-art segmentation methods in the AAPM thoracic auto-segmentation challenge [19]. Due to the limited patient cohort of 14 subjects, the leave-one-out cross-validation approach was applied to test the general performance of both image segmentation algorithms on multiple PMI datasets. The MA method is applied on all calculated PMI datasets (N=7). The 3D U-Net method is, due to practical restrictions (GPU calculation time), only applied to PMI datasets of the reference energy (70 keV) and the energy, which provided the best results for the MA method.

Organ-driven multi-atlas-based image segmentation

The in-house developed algorithm for MA based image segmentation consists of three major steps: the pre-processing, the deformation and the post-processing step. The atlas database is composed of atlases with its CT and OAR segmentation volume [20]. In the pre-processing step, the CT volume is denoised using the edge-preserving multi-threaded curvature flow image filter (5 iterations; time step = 0.05) implemented in the Insight Segmentation and Registration Toolkit (ITK).

In the deformation step, Elastix was used to apply a multi-stage deformable image registration between the CT volume of the unsegmented patient and the CT volumes stored in the atlas database [21]. This deformable image registration algorithm first calculates an affine transformation followed by a B-spline transformation. The calculated deformation field is then used to deform the CT and segmentation volume from all atlases within the atlas database to the unsegmented CT volume. For every OAR, a consensus between the deformed segmentation volumes is found by applying the simultaneous truth and performance level estimation (STAPLE) filter on the atlases with the highest normalized cross-correlation (NCC) coefficient around the OAR [22]. In the final post-processing step, the segmentations resulting from the STAPLE filter were morphologically smoothed to obtain the final segmentation. The method was processed on a 400 core HTCondor CPU cluster.

Two-step 3D U-Net deep learning method

The two-step 3D U-Net DL method is a 3D convolutional neural network architecture, which was applied twice in succession. First, the 3D bounding box location of each OAR was detected, and second, the OAR found in the previously detected bounding box was segmented.

As a result of the AAPM 2017 thoracic auto-contouring challenge, the original implementation of the deep learning (DL) algorithm, used in this study, was made available by Dr. Xue Feng from the University of Virginia. The algorithm was written

in Tensorflow (Python 2.7) and made use of the NVIDIA's CUDA® Deep Neural Network library (cuDNN) computational kernels.

To train the 3D DL convolutional neural network, a training dataset of atlases was created. Every atlas in the training dataset consisted of two 3D volumes: (i) the CT volume and (ii) the segmentation volume. Before learning the first step of the 3D U-net model, pre-processing of the CT and segmentation volume was performed. The segmentation volume contains information about the manual OAR segmentations wherein unique OAR flag IDs were assigned to every voxel in the volume. In this automatic pre-processing step, CT numbers between -500 HU and 1500 HU were normalized and both volumes (i and ii) were resized to unify voxel dimensions ($0.98 \times 0.98 \times 2.00 \text{ mm}^3$).

The first training step (900 epochs) was applied on down-sampled and cropped volumes. The normalized CT and segmentation volume were down-sampled to half of its original dimensions, where after the volume was cropped with 48 voxels to remove less important air voxels from both volumes. Unique OAR flag IDs were assigned in the segmentation volume, except for the eye lenses. No model was learned for the eye lens, because its volume was too small after down-sampling. Alternatively, the segmentation volume of the eye lens was morphologically subtracted from the segmentation volume of the eyes before down-sampling.

The network architecture contained three encoding and three decoding layers, used weighted cross entropy as loss function and the dropout was equal to 0.5. For each OAR, a bounding box with preset fixed sizes was determined. The bounding box size was equal to [80, 88, 88] voxels for the brainstem, [48, 88, 88] for the eyes and optic nerves, and [64, 120, 120] for the parotid glands. In the second step, one network (500 epochs) is trained per OAR to segment the foreground pixels. In the last automatic post-processing step, the segmented foreground pixels of the OAR were cleaned. The non-contiguous regions were removed, and the binary holes were filled using morphological operations. Because no model was trained for the eye lens, an algorithm was written to detect the binary hole in the eye segmentation automatically, and to identify these pixels as eye lens. In total, 28 neural networks were trained and applied on two graphics cards: the GeForce Titan Xp and the Quadro P6000.

Quantitative and qualitative segmentation accuracy assessment

For quantitative assessment of the image segmentation accuracy, three evaluation metrics were evaluated: the Dice similarity coefficient (DSC), the 95th percentile ($P_{95\text{th}}$) Hausdorff distance (HD), and the center of mass displacement (ΔCOM). Each metric compares the automatic segmentation S_{AUTO} and the manual delineation by the radiation oncologist S_{RO} , which serves as ground truth here. The DSC calculates the overlap between two 3D volumes and is equal to one for a perfect overlap and equal to zero without any overlap. The distance between two outer surfaces is

described by the HD, where the optimal outcome is equal to 0 mm and where increasing distances indicate less or no overlap [23]. To ignore the influence of a very small subset of inaccuracies in the automatically generated OAR segmentation, the 95th percentile of the HD is considered in this study. The Cartesian ΔCOM between two segmentations is a measure of the 3D position shift of the segmentation.

For qualitative assessment, two experienced radiation oncologists as well as an experienced radiation technologist individually scored the automatic image segmentation according to a discrete four-grade scale. The evaluation was done for all OARs on the PMI-70 dataset and additionally on the PMI-40 dataset for the optic nerves. A score of 1 was assigned to clinically unacceptable contours, for which it would take more time to modify than to restart a manual delineation. Scores 2 or 3 were assigned to clinically acceptable contours with major and minor changes, respectively. A score of 4 was assigned to clinically acceptable contours with none or negligible changes. This qualitative assessment was performed twice on all 14 patients.

Results

Quantitative assessment

The results of the three different evaluation metrics (DSC, HD, and ΔCOM) for the MA method applied on various PMI datasets are presented in [Figure 2](#).

As the PMI-70 dataset was originally used for manual contouring by the radiation oncologist, it was defined as reference PMI in the further analysis. The DSC between the manual and automatic contour generally was maximal when the MA method used PMI datasets of the lowest energy (i.e. PMI-40). The largest mean differences were found to be 2.6% for the parotid glands and 8.1% for the optic nerves. Smaller relative DSC differences were noticeable for the brainstem (0.4%) and the eyes (0.4%). In general, the HD values between manual and automatic segmentation reduced on average when the MA method was applied on the PMI-40 dataset compared to the PMI-70 dataset: 10.4% for brainstem, 4.2% for eyes, 20.6% for optic nerves and 11.9% for parotid glands. A similar trend was observable for the ΔCoM metric. The DSC for eye lenses were difficult to interpret due to the large discrepancies over the whole energy range of PMI datasets.

Considering the general improvements in the DSC, HD, and ΔCOM for the MA algorithm using the PMI-40 dataset, the 3D U-Net DL model was trained on the PMI-40 and PMI-70 datasets in a leave-one-out cross-validation approach. [Figure 3](#) depicts the three evaluation metrics for both segmentation methods using the PMI-40 and PMI-70 datasets.

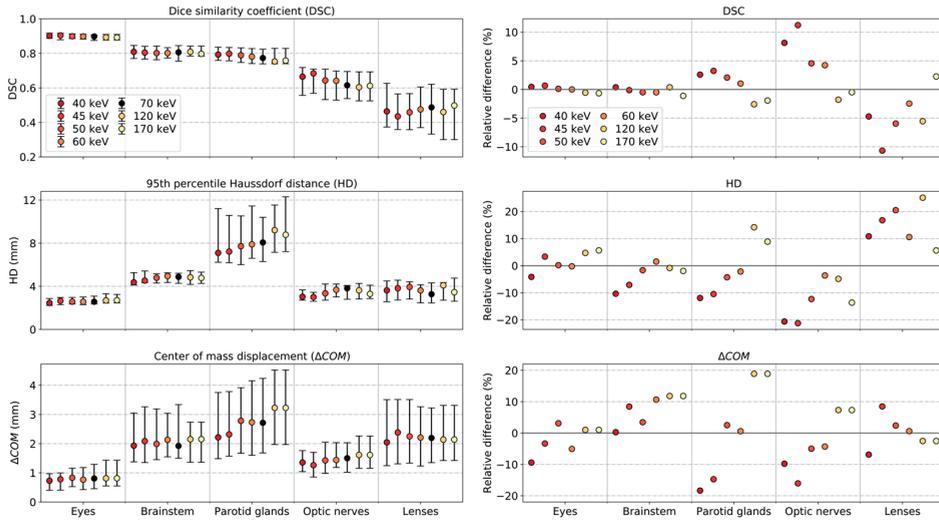


Figure 2. (a) Quantitative evaluation metrics calculated between the manual reference contour and the automatic segmentations generated by the multi-atlas-based image segmentation using pseudo-monoenergetic image (PMI) datasets of 7 different energies ranging from 40 keV to 170 keV. The markers indicate the median value, the whiskers represent the 25th and 75th percentile and the black marker is the reference energy (70 keV). (b) Relative differences between PMI datasets of different energies and the PMI of the reference energy (70 keV).

The DL neural networks trained on the PMI-40 dataset slightly underperformed compared to the one trained on the PMI-70 dataset. The DSC reduced on average with 4.0% for the brainstem and 4.1% for the optic nerves. Smaller relative DSC differences between the application on PMI-40 and PMI-70 datasets were noticeable for eyes (-0.2%) and parotid glands (-0.7%).

The DL approach outperforms the MA method on all three quantitative metrics for both PMI-40 and PMI-70 datasets, except for optic nerves. Compared to the DL method, the DSC for optic nerves was on average 28% and 19% larger for the MA method applied on PMI-40 and PMI-70 datasets, respectively.

Qualitative assessment

Comparing both segmentation approaches in a qualitative four-grade scoring system using the PMI-70 dataset, DL-based segmentations were less assigned to the ‘not clinically acceptable’ category, except for optic nerves (Figure 4). The application of the DL approach led to not clinically acceptable assignments (score 1) in less than 5% of all brainstem, eye lens and parotid gland segmentations. DL-based contours of the brainstem, eye, eye lenses and parotid glands were classified as acceptable with or without minor changes (scores 3 and 4) in 42.9%, 96.4%, 90.5% and 75.0%, respectively.

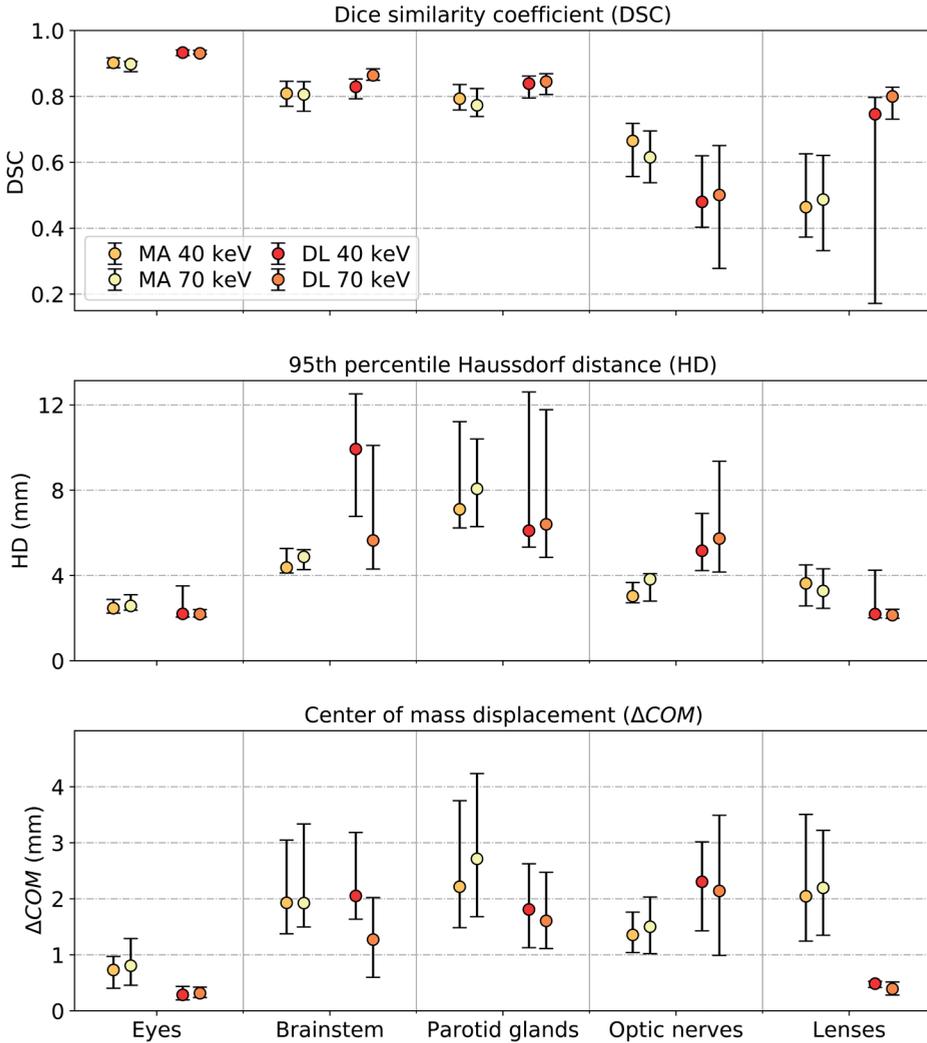


Figure 3. Quantitative evaluation metrics between the manual and automatic segmentations derived from pseudo-monoenergetic image (PMI) datasets of 40 keV and 70 keV for the multi-atlas (MA) and deep learning (DL) based image segmentation. The markers indicate the median value and the whiskers represent the 25th and 75th percentile.

In accordance with the quantitative assessment, the optic nerve segmentations by the MA method scored better than the DL method (Figure 5). Only 4.7% of the optic nerve segmentations using the MA method were assigned to the ‘not clinically acceptable’ category, where the DL approach resulted in 33.3%. The segmentations of the optic nerves using the PMI-40 datasets resulted in less assignments to the ‘major changes’ category and more assignments in the ‘minor changes’ category. The stacked bar charts of the three individual scorers are presented in the [Supplementary Material A](#).

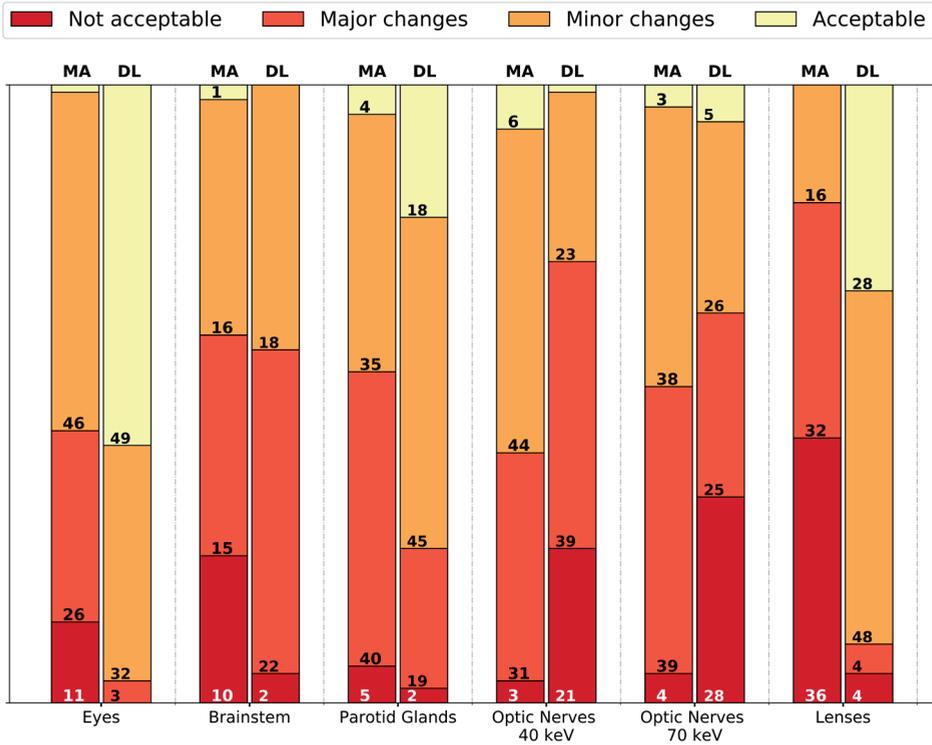


Figure 4. Stacked bar chart of the qualitative four-grade scoring (not clinically acceptable, clinically acceptable with major changes, clinically acceptable with minor changes, clinically acceptable) of the automatic multi-atlas (MA) and deep learning (DL) based image segmentations. The numbers in the bars indicate the occurrence in each category by the medical experts. The sum of the occurrence is equal to 84 for all organs (14 patients, 3 scorers and left/right), except for the brainstem (N=42).

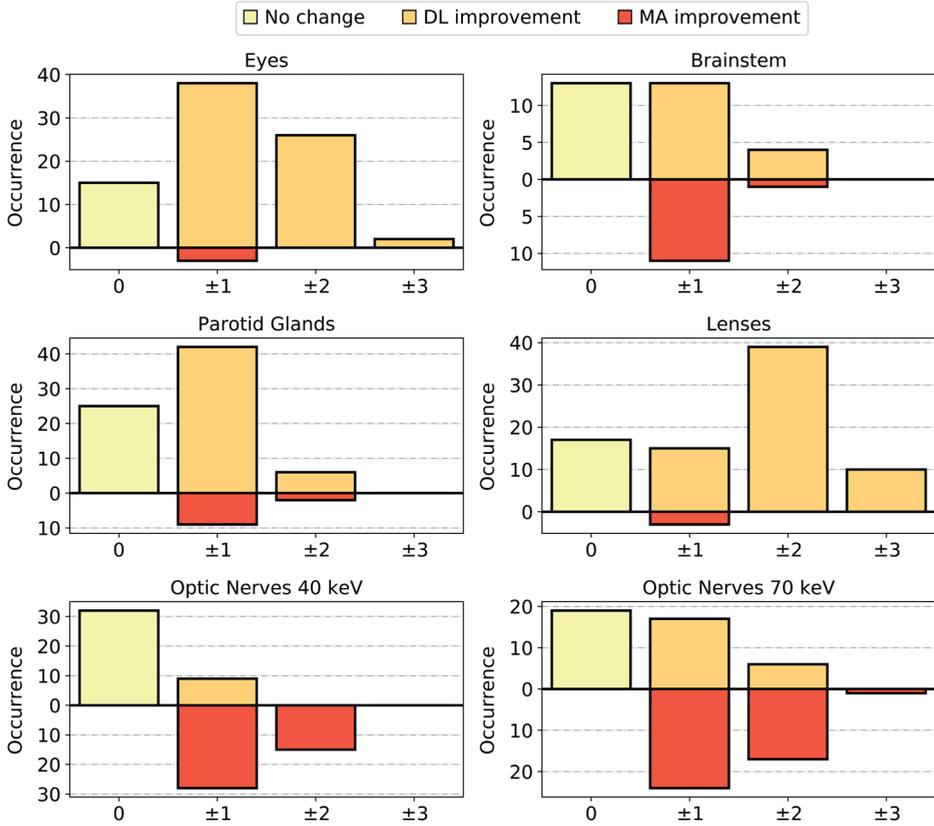


Figure 5. The relative occurrence of changes in the qualitative scoring between the multi-atlas (MA, red) and deep learning (orange) methods including all observers. If the scoring of both approaches was the same, it was categorized as not change (yellow). For the respective method, improvements of one to three qualitative scores were distinguished.

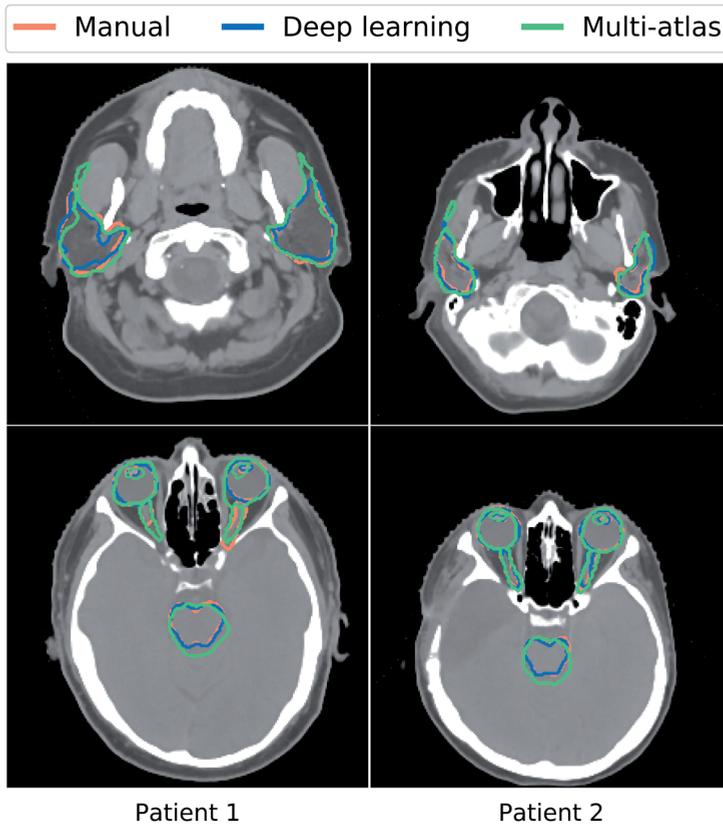


Figure 6. Comparison of the manual (orange), deep learning (DL; blue), and multi-atlas (MA; green) based image segmentation methods for all organs-at-risk for two patients.

Figure 6 gives a visual representation of the manual and automatic OAR delineation for both segmentation methods. In the left column, patient 1 with the overall lowest DSC is shown. The right column shows patient 2 with the overall largest DSC.

Calculation time

On average, the DL method demanded for training (53 ± 4) hours on a Geforce Titan Xp or Quadro P6000 (NVIDIA Corporation, Santa Clara, USA) for one neural network in our adopted leave-one-out cross validation. An average inference time of (20 ± 2) seconds was required to segment all OAR foreground pixels of one patient. The MA method which was performed on a 400 core HTCondor CPU cluster required around 10 minutes to segment all OAR foreground pixels of one patient.

Discussion

Much effort has been clinically made to study the advantages of DECT imaging in radiotherapy, to demonstrate its clinical relevance and accuracy in dose calculation especially for proton therapy as well as to implement DECT in a clinical workflow [16, 24-26]. However, to our knowledge, no study has been investigated automatic OAR segmentation methods on multiple PMIs, neither for MA methods nor DL approaches.

The MA and DL methods were quantitatively and qualitatively evaluated for various PMI datasets. Compared to the higher energies, for most OARs, the automatic segmentation using the MA method revealed better quantitative results using the PMI-40 dataset. This resulted in a DSC improvement of 8% for the optic nerves compared to the automatic segmentation generated using the PMI-70 dataset.

The PMI-70 dataset was also used for manual delineation. According to Wohlfahrt *et al.* 2017, this dataset reveals the best image-noise and contrast-noise ratio compared to the other six reconstructed PMIs [16]. Since this might have introduced a small bias in the results, the quantitative metric errors of the MA method were reported relatively to PMI-70.

Thereafter, the PMI-40 and PMI-70 datasets were used for neural network training. In the quantitative results of the 3D U-Net, the contrary was observed. Here, the segmentation results of the PMI-70 dataset were slightly better than the segmentation results using the PMI-40 dataset, which may be explained by the lower noise levels in the PMI-70 compared with the PMI-40 dataset. To quantify differences in image noise between PMI-40 and PMI-70, the CT numbers in Hounsfield Units (HU) within a uniform brain region were evaluated in a circular region-of-interest, which were (50 ± 9) HU and (43 ± 5) HU, respectively. Except for the cerebrospinal fluid, the brainstem is surrounded by brain tissues having nearly the same image intensity and therefore it is more difficult to perform a segmentation while having higher noise levels in the non-contrast cranial body region.

The DSC, HD, and the Δ COM metrics of the eye lenses were difficult to interpret due to the large discrepancies over the whole energy range. These discrepancies occurred because eye lenses are small and only consist of a few voxels. The inclusion or exclusion of a single voxel or only a few voxels will lead to a large change in the quantitative evaluation metrics.

Quantitatively and qualitatively, the 3D deep learning approach performed better than the multi-atlas method, except for the optic nerves. The optic nerves are a relatively small volume delineated on every single axial CT slice. Since the anatomical extension of optic nerves is not only in transversal direction, relatively small volumes on multiple axial CT slices do not necessarily form a connected 3D object. This very likely caused the increased difficulties to find a contiguous segmentation volume

of the 3D deep learning approach compared to the MA method. An overview of the organ volumes is listed in the [Supplementary Material B](#).

The quantitative performance of both automatic segmentation methods was compared to published results of other methods. Zaffino et al. [27] quantitatively assessed the performance of their automatic image segmentation software (PLASTI-MATCH) applied on neurological cancer patients. They reported (values extracted using the DataThief3 software) a median DSC of 0.79 for brainstem, 0.43 for optic nerves and 0.78 for parotid glands using a mean of 18 ± 3.5 atlases to segment each patient. Considering the median DSC, our MA method applied on the PMI-40 dataset performed similar with 0.81 for brainstem, 0.67 for optic nerves, and 0.79 for parotid glands. Our DL approach applied on the PMI-70 dataset performed better for some OARs, such as 0.86 for brainstem and 0.85 for parotid glands. Here, a lower median DSC of 0.50 was calculated for optic nerves. Our DSC obtained for parotid glands, brainstem, and eyes using the DL method were comparable with the reported inter-observer variability of 0.85, 0.83, and 0.83, respectively [28, 29].

Considering the fact that both approaches, MA and DL, performed well for some OARs, while poorly for others (e.g. the optic nerves), more effort needs to be put into learning and evaluating the algorithms on larger datasets. However, the median DSCs reported in this study were comparable with the reported inter-observer variability assessed on conventionally used 120 kVp CT images [3, 28, 29]. Both algorithms were already compared to state-of-the-art automatic contouring techniques in the ‘thoracic auto-segmentation challenge’ organized at the 2017 annual meeting of the American Association of Physicists in Medicine, in further studies they should also be evaluated on other body sites [19].

In terms of calculation time, the training time of the DL method is not the most important factor in OAR segmentation for radiotherapy, because the site-specific model training only needs to be performed once for a single DECT-derived PMI dataset of a specific energy and OAR combination. The time required to apply the automatic OAR segmentation on new patient datasets is more important in a clinical radiotherapy workflow, since the algorithm will be applied on each cancer patient being examined. The inference time of the DL method (± 20 seconds) is much faster than the ± 10 minutes of CPU cluster calculation time required by the MA method. Besides the improved OAR segmentation time of the DL method, the quantitative and the qualitative results were generally better for most of the OARs.

In this retrospective study, the automatic segmentations by state-of-the-art MA and DL approaches [19, 20] were evaluated between various PMI datasets (40 to 170 keV) on a limited atlas database of 14 primary brain-tumor patients that underwent DECT imaging to assess the influence of different image contrasts on delineation variability. Due to our adopted leave-one-out cross-validation approach, we do not expect a considerable change of the median value of the respective metric calcu-

lated for the MA method if more patients are included in this study. As a limitation of this retrospective study, the DL method was only trained on 13 patients (N-1), however, larger patient databases are desired for DL approaches.

Many automatic contouring algorithms in radiotherapy are still evaluated on conventional 120 kVp SECT images [19, 30]. As already demonstrated in previous studies [10, 16], PMI datasets can improve image quality in terms of contrast and noise as well as reduce beam hardening artifacts depending on the energy selected for PMI generation compared to conventional 120 kVp SECT. For other applications such as a non-automatic and dedicated bone mineral segmentation, it has also been shown that the use of PMI datasets could reduce inter- and intra-observer variability. However, to fulfill the requirements of the as-low-as-reasonably-achievable (ALARA) principle, only DECT scans were clinically acquired for the investigated patient cohort, since no added value for medical purposes was expected from a conventional 120 kVp SECT and an 80/140 kVp DECT scan in direct succession. Consequently, this retrospective study solely aimed at the evaluation of the performance of MA and DL methods on DECT-derived PMI datasets. After demonstrating the benefits of PMI datasets for automatic segmentation, a future prospective study could be designed to directly assess the differences between SECT and DECT considering the radiation protection law and an ethics approval.

Conclusion

For the first time, a quantitative and qualitative assessment of the multi-atlas and deep learning-based segmentation approaches was performed with DECT-based pseudo-monoenergetic images of different energies. For the multi-atlas method, PMIs calculated at lower energies generally resulted in better OAR segmentations. The neural network generally performed better than the multi-atlas method. However, the deep learning approach encountered difficulties with the higher noise levels in PMIs of low energy. Further studies on larger image datasets and other body regions is needed to compare the performance of both OAR segmentation algorithms between DECT-derived datasets and possibly conventional 120 kVp SECT scans.

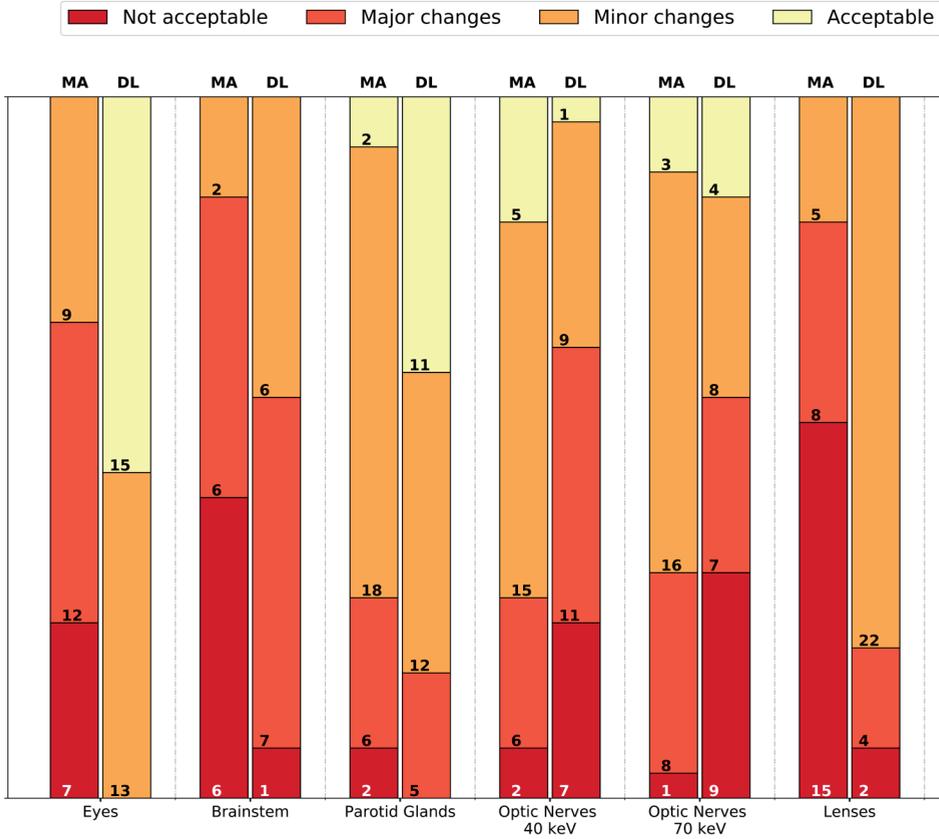
References

1. Ezzell, G.A., et al., *Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee*. Med Phys, 2003. **30**(8): p. 2089-115.
2. Mackie, T.R., et al., *Image guidance for precise conformal radiotherapy*. Int J Radiat Oncol Biol Phys, 2003. **56**(1): p. 89-105.
3. Sharp, G., et al., *Vision 20/20: perspectives on automated image segmentation for radiotherapy*. Med Phys, 2014. **41**(5): p. 050902.
4. Stapleford, L.J., et al., *Evaluation of automatic atlas-based lymph node segmentation for head-and-neck cancer*. Int J Radiat Oncol Biol Phys, 2010. **77**(3): p. 959-66.
5. Vinod, S.K., et al., *Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies*. Radiother Oncol, 2016. **121**(2): p. 169-179.
6. Yang, J., et al., *A statistical modeling approach for evaluating auto-segmentation methods for image-guided radiotherapy*. Comput Med Imaging Graph, 2012. **36**(6): p. 492-500.
7. Cheng, G., et al., *Multi-atlas-based Segmentation of the Parotid Glands of MR Images in Patients Following Head-and-neck Cancer Radiotherapy*. Proc SPIE Int Soc Opt Eng, 2013. **8670**.
8. Tong, N., et al., *Fully automatic multi-organ segmentation for head and neck cancer radiotherapy using shape representation model constrained fully convolutional neural networks*. Med Phys, 2018. **45**(10): p. 4558-4567.
9. Wohlfahrt, P., et al., *Clinical Feasibility of Single-Source Dual-spiral 4D Dual-Energy CT for Proton Treatment Planning Within the Thoracic Region*. Int J Radiat Oncol Biol Phys, 2018. **102**(4): p. 830-840.
10. Kuchenbecker, S., et al., *Dual energy CT: how well can pseudo-monochromatic imaging reduce metal artifacts?* Med Phys, 2015. **42**(2): p. 1023-36.
11. Yu, L., et al., *Virtual monochromatic imaging in dual-source dual-energy CT: radiation dose and image quality*. Med Phys, 2011. **38**(12): p. 6371-9.
12. Grant, K.L., et al., *Assessment of an advanced image-based technique to calculate virtual monoenergetic computed tomographic images from a dual-energy examination to improve contrast-to-noise ratio in examinations using iodinated contrast media*. Invest Radiol, 2014. **49**(9): p. 586-92.
13. Kaup, M., et al., *Dual-Energy Computed Tomography Virtual Monoenergetic Imaging of Lung Cancer: Assessment of Optimal Energy Levels*. J Comput Assist Tomogr, 2016. **40**(1): p. 80-5.
14. Michalak, G., et al., *Technical Note: Improved CT number stability across patient size using dual-energy CT virtual monoenergetic imaging*. Med Phys, 2016. **43**(1): p. 513.
15. Wichmann, J.L., et al., *Virtual monoenergetic dual-energy computed tomography: optimization of kiloelectron volt settings in head and neck cancer*. Invest Radiol, 2014. **49**(11): p. 735-41.
16. Wohlfahrt, P., et al., *Dual-energy CT based proton range prediction in head and pelvic tumor patients*. Radiother Oncol, 2017. **125**(3): p. 526-533.
17. Mangold, S., et al., *Single-source dual-energy computed tomography: use of monoenergetic extrapolation for a reduction of metal artifacts*. Invest Radiol, 2014. **49**(12): p. 788-93.
18. Wohlfahrt, P., et al., *Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monochromatic CT scans*. Int J Radiat Oncol Biol Phys, 2017. **97**(2): p. 427-434.

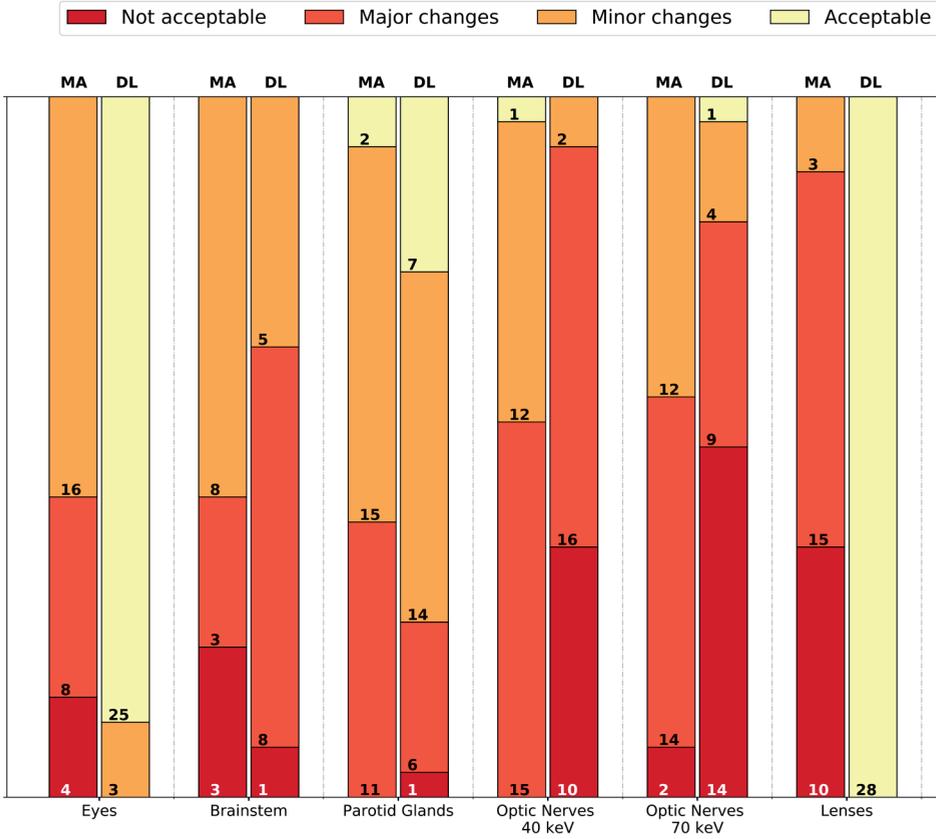
19. Yang, J., et al., *Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017*. Med Phys, 2018. **45**(10): p. 4568-4581.
20. van der Heyden, B., et al., *Automatic multiatlas based organ at risk segmentation in mice*. Br J Radiol, 2019. **92**(1095): p. 20180364.
21. Klein, S., et al., *elastix: a toolbox for intensity-based medical image registration*. IEEE Trans Med Imaging, 2010. **29**(1): p. 196-205.
22. Rohlfing, T., D.B. Russakoff, and C.R. Maurer, Jr., *Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation*. IEEE Trans Med Imaging, 2004. **23**(8): p. 983-94.
23. Huttenlocher, D.P., G.A. Klanderman, and W.J. Rucklidge, *Comparing Images Using the Hausdorff Distance*. IEEE transactions on pattern analysis and machine intelligence, 1993. **15**(0).
24. Bar, E., et al., *Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples*. Med Phys, 2018. **45**(1): p. 48-59.
25. Mohler, C., et al., *Experimental verification of stopping-power prediction from single- and dual-energy computed tomography in biological tissues*. Phys Med Biol, 2018. **63**(2): p. 025001.
26. Taasti, V.T., et al., *Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples*. Phys Med Biol, 2017. **63**(1): p. 015012.
27. Zaffino, P., et al., *Technical Note: plati-match mabs, an open source tool for automatic image segmentation*. Med Phys, 2016. **43**(9): p. 5155.
28. Deeley, M.A., et al., *Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study*. Phys Med Biol, 2011. **56**(14): p. 4557-77.
29. Faggiano, E., et al., *An automatic contour propagation method to follow parotid gland deformation during head-and-neck cancer tomotherapy*. Phys Med Biol, 2011. **56**(3): p. 775-91.
30. Raudaschl, P.F., et al., *Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015*. Med Phys, 2017. **44**(5): p. 2020-2036.

Supplementary Material

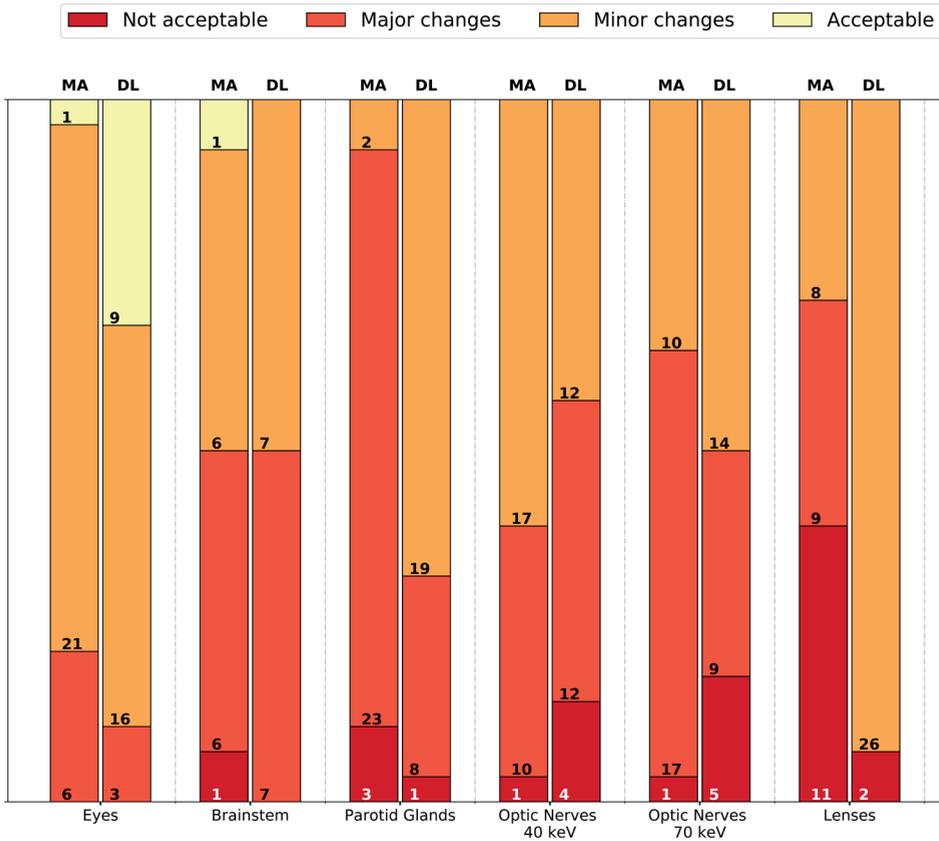
A. Individual observer scores



Supplementary 1. Qualitative scoring by the first medical specialist.



Supplementary 2. Qualitative scoring by the second medical specialist.



Supplementary 3. Qualitative scoring by the third medical specialist.

B. Organ-at-risk volumes

The average organ-at-risk (OAR) volumes were calculated for 14 patients. The volume is listed for five different segmentation methods: (i) manual segmentation, (ii) deep learning based segmentation on the 40 keV image dataset, (iii) deep learning based segmentation on the 70 keV image dataset, (iv) multi-atlas based segmentation on the 40 keV image dataset, and (v) multi-atlas based segmentation on the 70 keV data-set. The OAR volumes were determined considering the number of segmented voxels, the CT slice thickness, and the CT pixel spacing.

Supplementary 4. Average organ-at-risk (OAR) volumes (± 1 standard deviation) in mm³ of the manual segmentation. (Abbreviations: L, left; R, right; DL, deep learning; MA, multi-atlas)

In [mm ³]	Manual	DL 40	DL 70	MA 40	MA 70
Brainstem	23097 \pm 2647	20015 \pm 4684	22015 \pm 2757	26766 \pm 4105	26652 \pm 4756
Eye L	7864 \pm 650	7518 \pm 647	7742 \pm 633	8037 \pm 856	8068 \pm 1106
Eye R	8202 \pm 742	8013 \pm 670	8227 \pm 675	8393 \pm 755	8270 \pm 976
Lens L	183 \pm 30	166 \pm 85	132 \pm 57	217 \pm 77	240 \pm 113
Lens R	172 \pm 31	146 \pm 62	144 \pm 28	227 \pm 72	225 \pm 73
Optic Nerve L	524 \pm 159	295 \pm 155	295 \pm 233	665 \pm 122	681 \pm 134
Optic Nerve R	538 \pm 154	326 \pm 236	264 \pm 222	710 \pm 128	731 \pm 137
Parotid gland L	24040 \pm 10392	21534 \pm 9278	21301 \pm 8237	28427 \pm 12114	26974 \pm 10938
Parotid gland R	24629 \pm 10919	22349 \pm 9165	21982 \pm 8144	28820 \pm 10340	27999 \pm 9247

6

Chapter 6

Modelling of the focal spot intensity distribution and the off-focal spot radiation in kilovoltage X-ray tubes for imaging

Brent van der Heyden, Gabriel P Fonseca, Mark Podesta, Ivan Messner, Niklas Reisz, Ana Vaniqui, Heinz Deutschmann, Phil Steininger, and Frank Verhaegen

Published in

Physics in Medicine and Biology (PMB)

<https://doi.org/10.1088/1361-6560/ab6178>

Abstract

Background and Purpose: X-ray tubes for medical applications typically generate X-rays by accelerating electrons, emitted from a cathode, with an interelectrode electric field, towards an anode target. X-rays are not emitted from one point, but from an irregularly shaped area on the anode, the focal spot. Focal spot intensity distributions and off-focal radiation negatively affect the imaging spatial resolution and broadens the beam penumbra.

Materials and Methods: In this study, a Monte Carlo simulation model of an X-ray tube was developed to and evaluate the spectral and spatial characteristics of off-focal radiation for multiple photon energies. Slit camera measurements were used to determine the horizontal and vertical intensity profiles of the small and the large focal spot of a diagnostic X-ray tube. First, electron beamlet weighting factors were obtained via an iterative optimization method to represent both focal spot sizes. These weighting factors were then used to extract off-focal spot radiation characteristics for the small and large focal spot sizes at 80, 100, and 120 kV. Finally, 120 kV simulations of a steel sphere ($d=4\text{mm}$) were performed to investigate image blurring with a point source, the small focal spot, and the large focal spot.

Results: The magnitude of off-focal radiation strongly depends on the anode size and the electric field coverage, and only minimally on the tube potential and the primary focal spot size. In conclusion, an X-ray tube Monte Carlo simulation model was developed to simulate focal spot intensity distributions and to evaluate off-focal radiation characteristics at several energies.

Conclusion: This model can be further employed to investigate focal spot correction methods and improve cone-beam CT image quality.

Conflict of Interest

The authors disclosed receipt of joint financial support (Eurostars - MEIRS, reference number 11605; FFG, reference 861339; RVO, Rijksdienst voor Ondernemend Nederland) for the research of this thesis chapter.

Introduction

Diagnostic X-ray tubes are widely used for patient imaging in radiology and radiotherapy. In an X-ray tube, the primary electron beam, emitted from the cathode is accelerated by an interelectrode electric field and focused towards an anode target consisting of a high atomic number material such as molybdenum ($Z=42$) or tungsten ($Z=74$). The primary electron beam interacts with the target and produce photons by the deceleration of the electrons (i.e. Bremsstrahlung) or by characteristic X-rays [1]. Although photon emission is often assumed to originate from one single point source in Computed Tomography (CT) image reconstructions, this is a practical approximation and not strictly true. Photons are emitted from a small area on the surface of the anode, which is called the focal spot. Diagnostic X-ray tubes are commercially available with focal spot sizes typically ranging from 0.1 mm to 2.0 mm. A larger focal spot size negatively affects the spatial resolution of imaging and induces a wider X-ray beam penumbra [2-4].

After the first interaction round of the primary electron beam on the target anode, a fraction of the electron beam scatters back into the X-ray tube vacuum envelope and, in the presence of the interelectrode electric field, can be redirected back towards the target for a second round of interactions and can generate additional photon radiation [5]. Electrons can backscatter multiple times, yielding multiple generations of decreasing backscatter. A fraction of the re-entry electrons generates photons mostly outside the primary focal spot resulting in the so-called off-focal spot radiation. Off-focal spot radiation is also known as extra-focal radiation [6, 7] or off-focus radiation [8, 9] and its contribution cannot be neglected in X-ray tubes suitable for diagnostic imaging, as has long been indicated in the literature [6]. More recently, studies investigated the spectral characteristics of off-focal spot radiation in diagnostic X-ray tubes [10] and simulated the secondary radiation produced in transmission-type X-ray tubes [11].

The off-focal characteristics of X-ray spectra (position, energy, backscatter generation) are unmeasurable inside the vacuum envelope of the X-ray tube and are often ignored in X-ray tube simulation studies [1, 12-14]. In faster analytical codes for X-spectrum calculations such as SpekCalc [15-17] and Spektr [18] this effect is also not modelled. Currently there are no standardized methods to measure the off-focal component of X-ray spectra, and so Monte Carlo simulations provide an alternative to obtain these characteristics as simulated particles can be labeled according to their type and interaction history in user-specified categories.

The first goal of this study was to develop a realistic focal spot simulation model in an X-ray tube, and then to quantify the spectral characteristics and the spatial extent of the off-focal radiation due to the influence of the interelectrode electric field. This was investigated for an electron point source at 100 kV and two focal spot sizes at three commonly used diagnostic energies: 80, 100, and 120 kV. The effects of the enlarged focal spot size and the off-focal radiation on image quality was also investigated.

Materials and methods

Simulation model

A Monte Carlo simulation model of the ImagingRing system on rails (IRr) (medPhoton GmbH, Salzburg, Austria) was made using the Geant4 (v10.5) simulation toolkit [19]. The IRr system is a special type of cone-beam computed tomography (CBCT) scanner that allows non-isocentric CBCT acquisitions using a monobloc system (X-ray tube and generator) and a flat panel detector that rotate independently to optimize the field of view at the detector level. Geometrical and technical specifications of the IRr system and its components were provided by the IRr manufacturer (medPhoton).

The customized monobloc system (HF1 R-23 ORB X20P, IMD Generators, Grasso, Italy) consists of a diagnostic X-ray tube with a rotating disk-shaped (mesh volume) tungsten-rhenium target anode (95% W, 5% Re) that has an anode angle of 10 degrees and a diameter of 6.4 cm, an interelectrode distance of approximately 1.5 cm, and an inherent tube filtration of 1.3 mm aluminium at the beam exit window. A discretized interelectrode electric field was simulated perpendicular to the anode central axis in discrete and uniform electric field slabs (50 bins), each with their own electric field strength (V/m), over the active area of the anode (Figure 2b). Due to the anode angle of 10 degrees, the interelectrode distance is different for every modelled field slab. To compensate for this, the nominal electric field strength assigned to each discretized electric field slab (Figure 2b) is corrected by a factor ranging between 0.87 (E_1) and 1.17 (E_2). Ignoring this variation in electric field strength (V/m) between the slabs would result in X-ray spectra with incorrect maximum energy. To illustrate, this would mean that when an electron is emitted in e.g. E_1 (Figure 2b), without nominal field strength correction, the maximum energy in the X-ray spectrum will be 13% too high. Due to this correction factor, the electric field slabs closer to the central axis of the anode (E_2 in Figure 2b) will have a higher electric field strength than the electric field slabs further from the central anode axis (E_1 in Figure 2b).

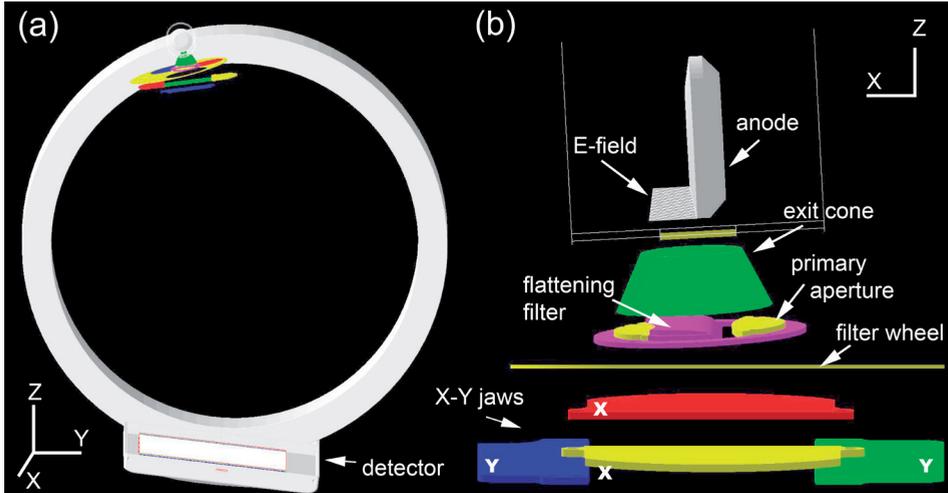


Figure 1. (a) Full Monte Carlo (Geant4) model of the Imaging Ring r (IRr) system and (b) the Monobloc model in more detail.

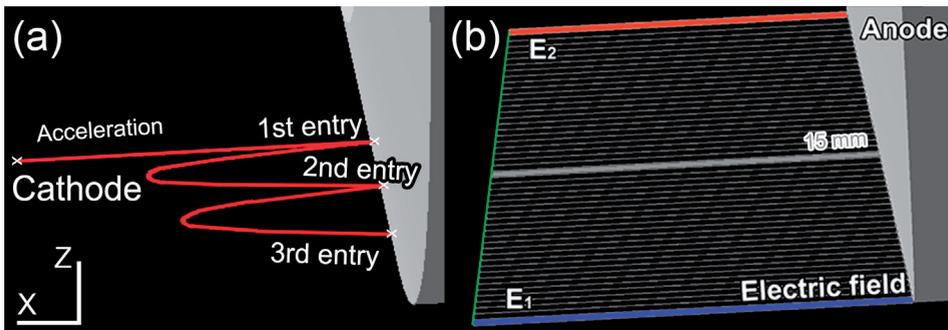


Figure 2. In the left panel (a), a sketch of the backscatter and the generations of re-entry electrons for an accelerated electron between cathode and anode. The inter-electrode field between the cathode and the anode is present, but not visualized in this sketch. In this example, the primary emitted electron first hits the target (1st entry), backscatters into the vacuum envelop, bends back towards the anode due to the influence of the electric field (2nd entry), and repeats the same interaction processes until the electron is stopped inside the anode (3rd entry). In the right panel (b), the simulated electric field slabs are visualized between the cathode (green line) and the anode.

Local electric field distortions causing misfocussed electrons and the individual electric field lines ending perpendicular on the target surface are ignored in this simulation model. The spatial extent of the electric field in the Y-direction (14 mm) and the Z-direction (13mm) was determined by manual iteration after the focal measurements until the electric field was able to accelerate all electrons from the optimized small and large focal spot cathode emission fluence.

Focal spot

Focal spot measurements

In contrast to what is often assumed in CBCT reconstructions, the X-rays generated in the X-ray tube are not originating from one single spot at or within the target. Although the manufacturers aim to focus the electron beam towards as small a spot as possible with the electric field, the focal spot still has a finite spatial distribution which is often not documented. At best, a square or rectangular spot size is quoted which is a first order approximation to the real focal spot intensity distribution. Focal spot measurements are required to have realistic focal spot intensity distributions which can then be used in the Monte Carlo model of the IRr monobloc system. Horizontal and vertical measurement profiles were provided by the manufacturer for the small and the large focal spot, and were acquired with a slit camera (L659117, PTW, Freiburg, Germany) and an XRD 1642 Perkin Elmer flat-panel detector (40.96 x 40.96 cm², 1024 x 1024 px²) [20]. The slit camera consisted of a lead disc with a 0.7 cm opening which was positioned 12.4 cm away from the focal spot.

To reduce noise effects in the next steps of the focal spot shape optimization, the horizontal and vertical profile shapes from the experiment were fitted in 1D to the sum of maximum four normal distributions wherein each distribution was characterized by three parameters: mean, standard deviation, and intensity. Noise effects were reduced to limit the number of beamlets required for the focal spot optimization. After the determination of the fit parameters, a 2D image of the small and large focal spot was obtained by convolving the fitted horizontal and vertical slit camera profiles (Figure 3).

Based on the focal spot measurements that showed a non-rectangular distribution and included off-focal radiation, the sizes of the small and the large focal spot were determined in the horizontal and the vertical profiles of the slit experiments at 15% of the maximum intensity to distinguish signal from noise. The small focal spot was 0.66 mm in the horizontal direction and 0.64 mm in the vertical direction, the large focal spot was 0.90 mm in the horizontal and 1.07 mm in the vertical direction.

Focal spot simulations

An optimization model was developed to determine the realistic cathode electron emission fluence in the IRr X-ray tube model, including the simulation of the electric field. In this optimization process, the cathode electron emission fluence was discretized in a regular rectangular grid (green plane in Figure 2b) with 21 horizontal grid points between -3 mm and + 3 mm, and 33 vertical grid points between -8 mm and +8 mm. The boundaries of this rectangular grid were aligned with a 6x6 mm² imaging beamlet detector that was simulated with a 60x60 μm² resolution at 12.4 cm dis-

tance from the focal spot representing the experimental slit camera measurements provided by the IRr manufacturer. Only photons with an incidence angle smaller than 0.3 degrees from the perpendicular X-ray beam axis were detected to mimic the slit camera experiment. A total of 693 (=21x33) emitted electron beamlets were simulated with 50E6 histories per electron beam and a bremsstrahlung splitting factor of 20,000 as variance reduction technique to reduce calculation time.

After the simulation of the 693 two-dimensional electron beamlet projections ($P_{1,...,n}$), their weighting factors ($w_{1,...,n}$) were optimized according to the cost function described by Equation 1 to match the convolved 2D distribution of the small and the large focal spot separately (Figure 3c and Figure 3d). The least-square nonlinear curve-fitting function 'lsqcurvefit' of the optimization toolbox in MATLAB® R2018b (The MathWorks Inc, Natick, MA) was used to minimize this cost function. During the optimization process, the minimum and maximum value of all weight factors were restricted to 0 and 1, and as constraint the sum of all weight factors must be equal to 1.

$$FS_{conv} = w_1 \cdot P_1 + w_2 \cdot P_2 + .. + w_n \cdot P_n \quad (\text{Eq. 1})$$

Electron backscatter simulations

A class was added to the Geant4 simulation to score particle characteristics in the X-ray tube such as the kinetic energy and the interaction position for every particle generation (Figure 2a). When an electron backscatters for the first time after exiting the target, it belongs to the first generation of backscattered electrons. An electron belongs to the second generation when an electron of the first generation re-enters the target, i.e. it has entered the target three times. The same terminology is applied on subsequent generations.

The derived electron emission fluence distributions of both focal spot sizes were used in the simulations of three different X-ray tube potentials: 80, 100, and 120 kV. Additionally, an electron point source for 100 kV was simulated for comparison with Ali and Rogers [10]. The added Geant4 class also allows for the calculation of the backscatter ratio, re-entry ratio, and bremsstrahlung production. 12.5E9 histories were simulated for every tube potential and focal spot combination.

Steel sphere simulation

A steel sphere ($\rho=7.7 \text{ g/cm}^3$) of 4mm diameter was simulated in the IRr model geometry and projections were scored using $0.4 \times 0.4 \text{ mm}^2$ resolution using an X-ray tube potential of 120 kV for three source configurations including off-focal spot radiation: (i) one electron beamlet as point source, (ii) the small focal spot, and (iii) the large focal spot. The source-to-object distance was 63.1 cm and the object-to-detector distance was 52.6 cm (IRr specifications). First, horizontal and vertical cross-section

profiles through the X-ray projections were evaluated and the full width at half maximum (FWHM) was calculated for each profile. Next, the circular edge spatial frequency (ESF) functions were calculated according to Baer [21], and then the modulation transfer functions (MTFs) were determined for all source configurations. 5E9 million electron histories were simulated considering the shape of the source configuration and bremsstrahlung splitting (x20,000) was applied as variance reduction technique.

Results

Focal spot measurements

The horizontal and vertical nominal profiles of the small focal spot (Figure 3a) and the large focal spot (Figure 3b) were used to convolve the small focal spot (Figure 3c) and the large focal spot (Figure 3d) at the level of the slit camera (12.4 cm from the focal spot).

Focal spot simulation

From the 693 simulated electron beamlets, 168 beams did not produce a signal at the pinhole camera detector (weight equal to 0) which means that some beamlets in the rectangular electron emission grid points did not produce any signal on the simulated 6x6 mm² detector. The summed distribution of the individual simulated beamlet grid points that produced a signal is shown in Figure 4a before optimization. After optimization, the individual beamlets were optimized according to Equation 1 to obtain the small focal spot (Figure 4b) and the large focal spot (Figure 4c) electron grid point weighting factors which were used in the electron backscatter simulations.

Electron backscatter simulations

The electron and photon statistics were first determined for a single electron beamlet at 100 kV and then for three different X-ray spectra (80, 100, and 120 kV) for the small and the large focal spot. In Table 1, the electron backscatter ratio (in %), report the amount of backscattered electrons from the anode compared to the first entry, and the electron entry ratio (in %) reports the number of re-entering electrons compared to the first entry. The first entry ratio and the 3rd generation bremsstrahlung are not listed in Table 1 because the first entry ratio is equal to 100% for all energies and the 3th generation bremsstrahlung was below 0.01%. Most of the re-entry electrons hit the target outside the focal spot. For the small focal spot simulation at 120 kV, only 1.3% of the (9.6%) 2nd entry and 2.1% of the (1.7%) 3rd entry electrons hit the target inside the high focal spot intensity distribution. For the large focal spot at

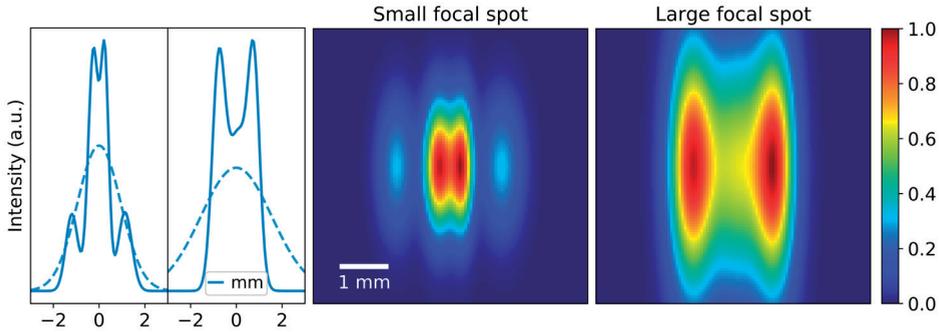


Figure 3. Focal spot slit measurements at the level of the slit camera at 12.4 cm from the focal spot in the horizontal (solid line) and vertical (dashed line) direction through the center of the focal spot for (a) the small focal spot and (b) the large focal spot. The normalized images of the convolved small and large focal spot profiles are shown in (c) and (d).

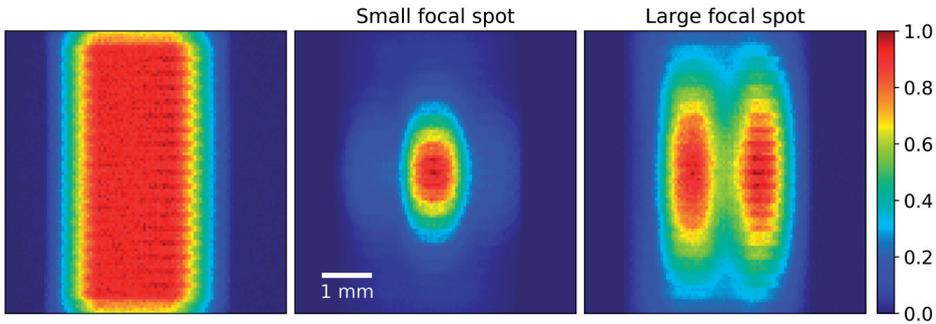


Figure 4. (a) The summed distribution of all individual beamlets before focal spot optimization where all the initial electron beamlets have equal weight. After individual beamlet optimization (Equation 1), the beamlet weights are optimized to obtain the focal spot image of the small focal spot (b) and the large focal spot (c). All panels are normalized to the maximum intensity between 0 and 1.

120 kV, these percentages were equal to 1.4% for the 2nd entry and 2.2% for the 3rd entry electrons. The electron and photon statistics were also simulated for a 20-degree anode angle instead of a 10-degree anode angle and are added as [Supplementary Material](#).

Figure 5 presents the calculated energy spectra at all angles for the different generations of backscattered electrons, entry electrons, and produced X-rays inside the X-ray tube target for the small and the large focal spot. The values on the vertical axis indicate the occurrences in every energy bin (in %) relative to the number of 1st entry electrons. The ‘fourth generation+’ in the bottom row also contains all higher generations.

Table 1. Small and large focal spot backscatter, entry, and X-ray production ratios for different generations and three X-ray tube potentials (80, 100, and 120 kV) for a 10-degree anode angle. The ratios of the simulated beamlet at 100 kV are also listed.

Backscatter ratio (% of the 1 st entry)	80 kV		100 kV		120 kV		
	<i>Small</i>	<i>Large</i>	<i>Beamlet</i>	<i>Small</i>	<i>Large</i>	<i>Small</i>	<i>Large</i>
1 st generation	47.9		48.3			48.4	
2 nd generation	6.0	5.8	6.5	6.3	6.0	6.5	6.2
Entry ratio (% of the 1st entry)							
2 nd entry	9.0	8.7	9.5	9.3	9.0	9.6	9.2
3 rd entry	1.5	1.4	1.6	1.6	1.5	1.7	1.6
X-ray production ratio (% of the 1st entry)							
1 st generation	1.11		1.67			2.31	
2 nd generation	0.02		0.03			0.05	

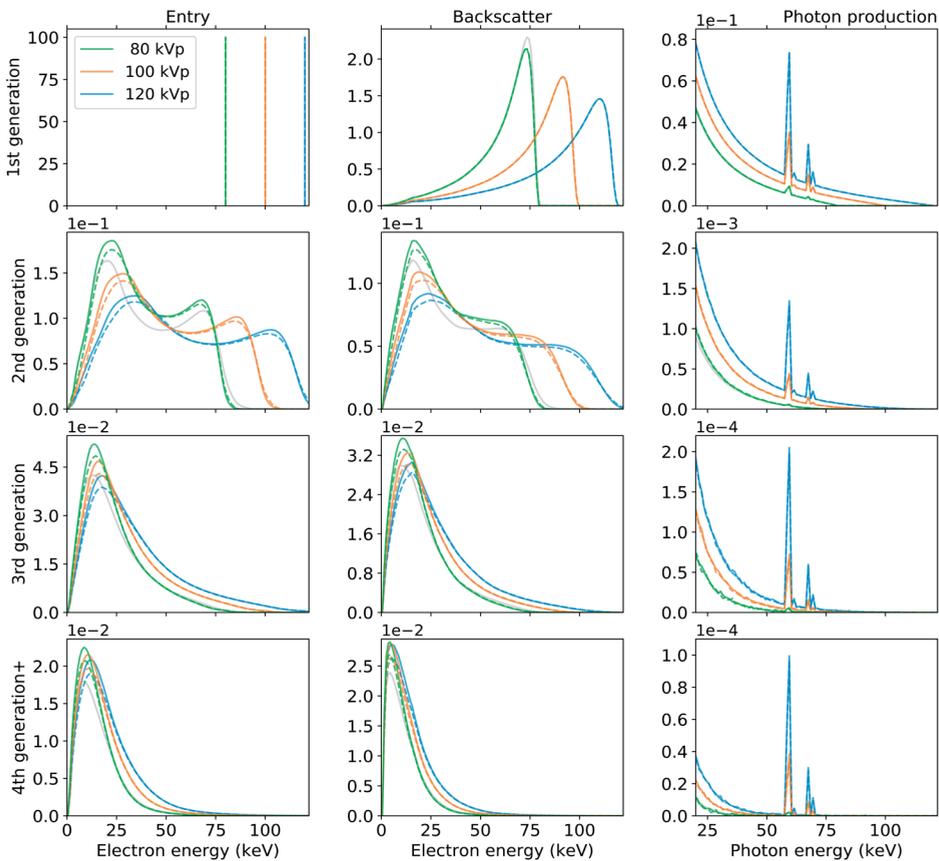


Figure 5. Calculated energy spectra of the entering electrons (left column), back-scattered electrons (middle column), and the X-ray production at all angles (right column) on a 10-degree anode angle for the small focal spot (solid lines) and the large focal spot (dashed lines). The energy spectra of a simulated small focal spot at 80 kV for a 20-degree anode angle is also plotted as a solid grey line.

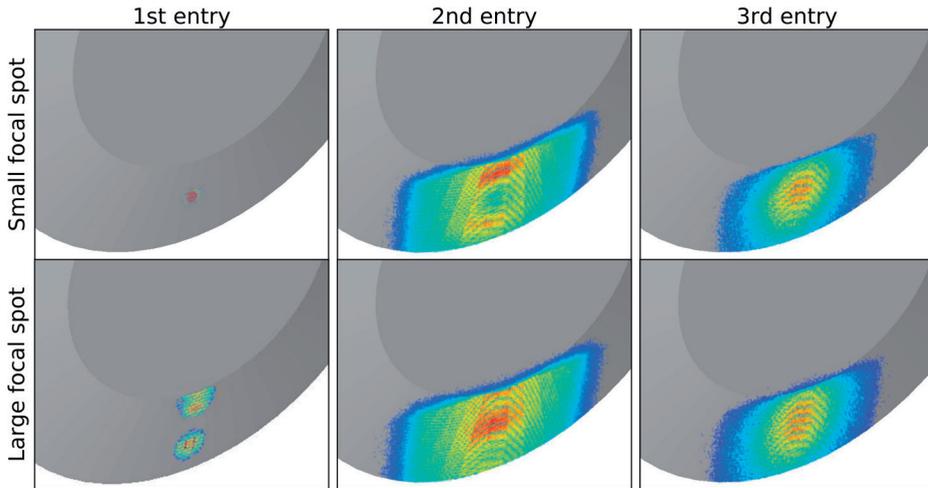


Figure 6. Spatial distribution of the entry electrons for the small (upper row) and the large (bottom row) focal spot for 120 kV. The frequency distribution of each entry generation is normalized to the maximum number of hits per entry.

In **Figure 6**, the three-dimensional frequency distribution of the entering electrons is visualized. The colormap is normalized to the maximum number of hits in the frequency distribution per generation and only the higher frequency bins are visualized. Therefore, no electron hits are plotted on a large part of the anode surface. The colormap of the frequency distributions in the 2nd and 3rd entry plots of **Figure 6** are optimized for visualization. Here, the magnitude of the frequency distributions of the 2nd and 3rd entry is approximately 5200 and 15,800 times smaller than the 1st entry for the small focal spot and is approximately 1600 and 5000 times smaller than the 1st entry for the large focal spot.

The normalized intensity distributions of the focal and off-focal radiation calculated at 120 kV for the small and the large focal spot are depicted in **Figure 7**. The focal distribution has discrete levels due to the primary electron fluence sampling. Summed distributions are calculated for the Y-dimension and the Z-dimension. In **Figure 7**, the electric field boundaries adopted in this X-ray tube simulation ($H=1.3$ cm/ $W=1.4$ cm) are indicated by the vertical gray lines. Electrons also re-enter the anode outside the electric field because higher generation electrons are not spatially limited by the simulated electric field. Off-focal radiation can be produced outside the electric field because backscattered electrons were sufficiently deflected by the electric field before escaping the field (negative X-momentum, **Figure 2**).

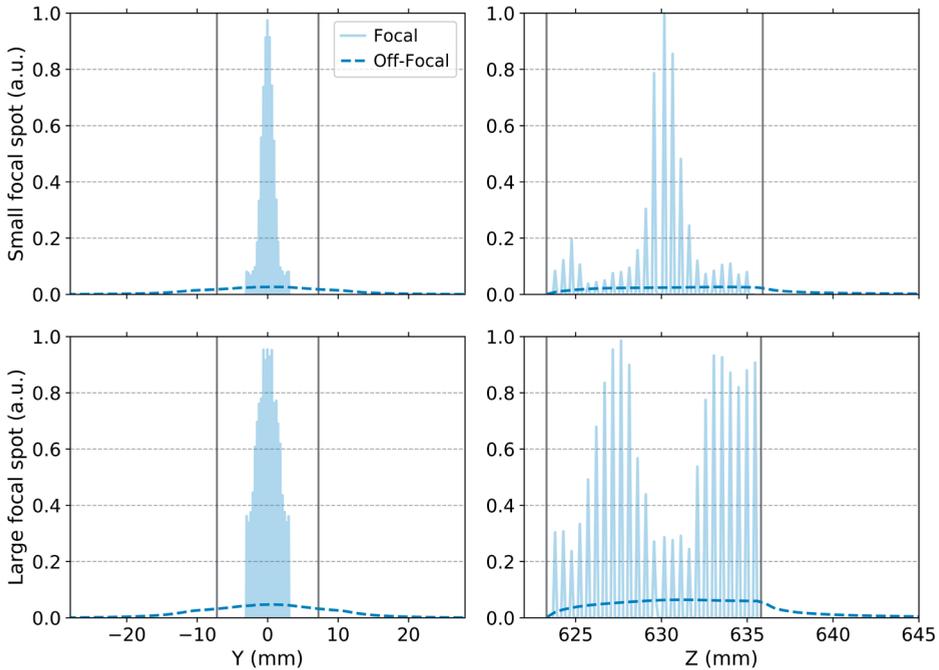


Figure 7. The simulated and summed intensity distribution of the focal (solid) and off-focal (dashed) radiation at 120 kV in the IRr X-ray tube for 1 dimension (Y or Z, in mm). Intensity distributions are shown for the small and the large focal spot. The vertical grey lines indicate the size of the electric field used in the simulation.

Steel sphere simulation

An imaging simulation of a small steel sphere (4 mm diameter) was performed for three different source configurations to evaluate the blurring effect at projection level due to the finite focal spot intensity distributions (Figure 8). The FWHMs of the horizontal and vertical cross-profiles were 6.94 and 6.91 mm for the ideal point source, 5.75 and 6.02 mm for the small focal spot, and 4.98 and 5.38 mm for the large focal spot configuration. The circular ESF functions show the steepest curve for the point source simulation and the flattest curve for the large focal spot. The 10% MTF values were 1.82 mm^{-1} for the ideal point source, 0.95 mm^{-1} for the small focal spot, and 0.69 mm^{-1} for the large focal spot.

Discussion

A detailed Geant4 Monte Carlo simulation model was developed to simulate X-ray tube focal spots and evaluate the characteristics of non-focal spot radiation at different generations under the influence of different interelectrode electric field strengths.

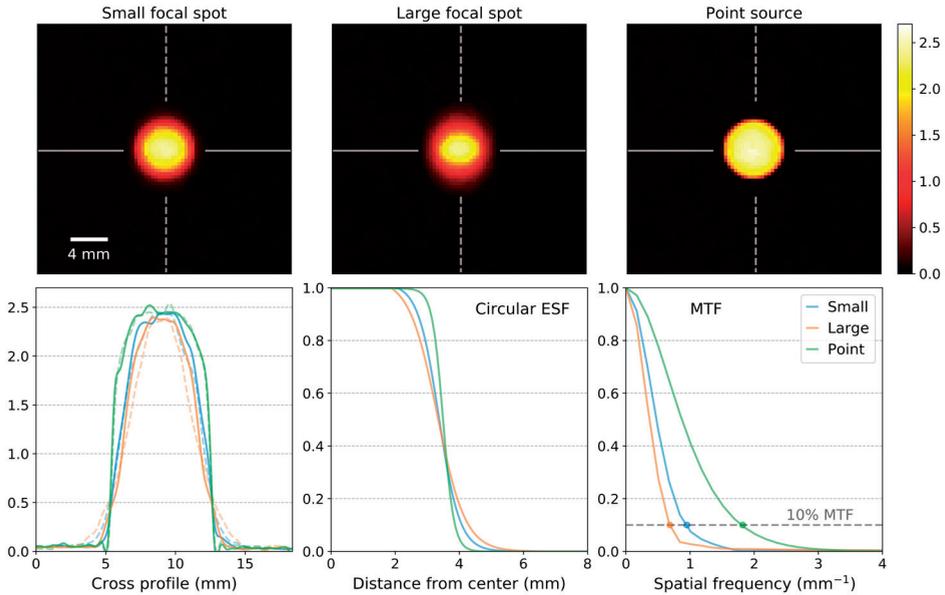


Figure 8. The top row visualizes the simulated projections of a steel sphere with three source configurations at 120 kV for the small focal spot (blue), the large focal spot (orange), and a point source (green). Horizontal (solid) and vertical (dashed) cross profiles were obtained through the image centers indicated by the white cross-hair. The bottom row shows both cross profiles, the circular edge spatial frequency (ESF) functions and the modulation transfer functions (MTFs) for all source configurations.

With the slit camera experiments provided by the manufacturer, the shape of the focal spot was reconstructed, and the electron emission fluence was determined for the small and large focal spot by solving an optimization function iteratively. The simulation model published by Ali and Rogers [10] was the most advanced model that was available in the literature. In Ali and Rogers [10], the off-focal component was simulated with an electron point source for different types of X-ray tubes and target materials using a uniform perpendicular electric field; e.g. mammography 26 kV (Mo), diagnostic 100 kV (W), and orthovoltage 250 kV (W) X-ray tubes. The model described in this study is capable to simulate electric emission fluences inside the X-ray tube, e.g. for the small and the large focal spot, and operates at different electric field strengths using an improved electric field simulation model that subdivides the electric field in slabs with corrected electric field strengths. Although only tungsten target materials are investigated, the model described in this study is also configurable for other common target materials and energy ranges such as molybdenum and rhodium in mammography between 22 and 40 kV [22, 23].

Ali and Rogers [10] also reported electron backscatter coefficients or ratios (in %) for the diagnostic 100 kV X-ray tube (W) scenario simulated with an electron beam-

let. Although backscatter ratios are difficult to compare between target geometries, because of its anode angle dependency, similar first-generation backscatter ratios are reported for our diagnostic 100 kV X-ray tube electron beamlet simulation (48.3%) and Ali and Rogers [10] (50.6%). Higher generation of backscatter ratios depend on the number of re-entry electrons and thus on the size and strength of the electric field combined with the size of the target anode. The lower generation electron backscatter ratios reported in this study are explained by a smaller electric field size laterally, and smaller target size $H=1.3$ cm/ $W=1.4$ cm compared to the much larger $H=2.5$ cm/ $W=10$ cm in Ali and Rogers [10] which could be too wide for a realistic diagnostic imaging X-ray tube considering that the electric field forms an electrostatic lens to obtain relatively small focal spot sizes. Smaller targets require a smaller volume wherein the E-field extends, and thus, less electrons will re-enter the target, backscatter again, or produce characteristic X-rays or bremsstrahlung. This is because electrons will exit the electric field lines and will not be bent towards to target again. In the manufacturing process of X-ray tubes, limiting the lateral extent of the electric field will reduce the off-focal radiation component in the imaging beam.

For this X-ray tube setup (i.e. interelectron distance and target size) the effect of the non-focal spot irradiation is limited for diagnostic imaging purposes such as cone beam CT imaging. In the model of a conventional diagnostic 120 kV X-ray spectrum, 2.3% of the primary electrons (1st entry) produces characteristic X-rays or bremsstrahlung and only 0.05% of the primary electrons will produce higher generation X-rays which have a much softer bremsstrahlung spectrum. This increases the probability that the extra-focal spot irradiation will be filtered by either the inherent tube filtration, by the added spectral filters used for diagnostic imaging, or by the X-ray beam collimators (i.e. X-Y jaws).

The anode angle has a small effect on the backscatter ratios and on the entry ratios listed in Table 1 (10-degrees) and as Supplementary Material (20-degrees). For the 80 kV focal spot simulations, the 1st generation backscatter ratio was 1.5% lower for the 10-degree anode angle compared to the 20-degree anode angle. The energy spectrum of the backscattered electrons simulated for a 20-degree anode angle have a larger high energy peak, which caused a lower 2nd generation entry ratio of approximately 1%, a lower 2nd generation backscatter ratio of 0.5%, and thus a lower 3rd entry ratio of 0.2%. No differences in the X-ray production ratios were observed between the two anode angles because the heel effect affects the X-ray spectrum only after the photon production.

This study focused on the off-focal radiation effect caused by the re-entering of backscattered electrons in the anode material. However, the contaminating extra-focal spot irradiation that softens the X-ray spectrum and broadens the angular distribution of the imaging beam is not only originating from the electron re-entry

effect inside the X-ray tube vacuum. Also, Compton and Rayleigh scatter interactions will occur in the anode material, the beam exit window, the additional spectral filtration that absorbs the low energy photons that only contribute to the patient skin dose, and in the imaging beam collimating aperture. Furthermore, backscattered electrons deflected under the influence of the electric field could also produce bremsstrahlung in the X-ray tube shielding made of high Z materials. Therefore, the global contribution of extra-focal radiation will be larger than the values reported in this study which are only due the electron re-entering effect on the target.

The focal spot size and its irregular intensity distribution also needs careful investigation for other applications such as precision image-guided animal radiotherapy. Here, collimators with relatively small exit windows, ranging between 0.5 mm and 10 mm, are used to collimate the orthovoltage treatment beam. Especially for the smallest exit windows (≈ 0.5 -1 mm), the beam collimator serves as a pinhole camera and projects the focal spot intensity distribution on the animal. Therefore, lacking focal spot intensity models in preclinical radiotherapy limits the accuracy of dose calculations and complicates the modelling of precision irradiation beams and their penumbrae [24, 25]. Exact modelling of focal and off-focal spot could improve the preclinical dose calculations in precision radiotherapy for the smallest beam sizes.

For imaging purposes, efforts should be made, and hardware solutions should be investigated to limit the focal spot size and limit the contribution of extra-focal spot irradiation as these effects lead to an increased imaging beam penumbra and leads to a degraded image quality (Figure 8).

In future studies, parameters of this simulation model such as the electric field strength and size, the target material and target mesh volume, and the electron emission fluence can easily be modified to investigate the adverse effects of focal spot distributions and extra-focal spot irradiation in X-ray projection space, or in image space after the (CB)CT reconstruction, and even on the patient CBCT imaging dose [26, 27] which is of interest when CBCT imaging is used in many fractions over the whole radiotherapy course for example. This model could be used to explore projection-based correction methods for the focal spot induced blurring, such as projection-based deconvolution methods using point-spread functions [28, 29] or artificial intelligence [30].

Conclusion

This study described the development of a diagnostic X-ray tube, Monte Carlo simulation model which includes the electric emission fluence intensity distribution for two focal spot sizes and the interelectrode electric field to simulate the focal spot intensity distribution and off-focal spot radiation. For our X-ray tube setup for CBCT imaging, the developed simulation model can be used to investigate the adverse effect of (off-)focal spot radiation on image quality.

References

1. Verhaegen, F., et al., *Monte Carlo modeling of radiotherapy kV x-ray units*. Physics in Medicine and Biology, 1999. **44**: p. 1767-1789.
2. Law, J., *The influence of focal spot size on image resolution and test phantom scores in mammography*. Br J Radiol, 1993. **66**: p. 441-446.
3. Ouandji, F., et al., *Impact of focal spot size on the spatial resolution of a digital x-ray imaging system for small animal studies*. Biomedical Diagnostic, Guidance, and Surgical-Assist Systems IV, 2002. **4615**.
4. Chen, L., et al., *Spatial resolution properties in cone beam CT: a simulation study*. Med Phys, 2008. **35**(2): p. 724-34.
5. Kim, S.H., et al., *Validation Test of Geant4 Simulation of Electron Backscattering*. IEEE Transactions on Nuclear Science, 2015. **62**(2): p. 451-479.
6. Birch, R., *The spectrum and intensity of extra-focal (off-focus) radiation*. Br J Radiol, 1976. **49**: p. 951-955.
7. Bootsma, G.J., et al., *Monte Carlo analysis of beam blocking grid design parameters: Scatter estimation and the importance of electron backscatter*. Med Phys, 2018. **45**(3): p. 1059-1070.
8. Thoræus, R., *The Amount of Off-focus Radiation in Beams from Various Types of Roentgen Tubes*. Acta Radiologica, 1937. **18**(5): p. 753-760.
9. Roeck, W.W., J.T. Martin, and R. Beach, *Analysis of contributing factors to the occurrence of off-focus radiation (OFR)*. Med Phys, 1992. **19**(5): p. 1201-4.
10. Ali, E.S. and D.W. Rogers, *Quantifying the effect of off-focal radiation on the output of kilovoltage x-ray systems*. Med Phys, 2008. **35**(9): p. 4149-60.
11. Boone, M.N., et al., *Secondary radiation in transmission-type X-ray tubes: Simulation, practical issues and solution in the context of X-ray microtomography*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2012. **661**(1): p. 7-12.
12. Bazalova, M. and F. Verhaegen, *Monte Carlo simulation of a computed tomography x-ray tube*. Physics in Medicine and Biology, 2007. **52**(19): p. 5945-5955.
13. Jia, P., Y. Xie, and S. Bao, *Monte Carlo simulation of X-ray tube spectra with PENELOPE*. Congress on Medical Physics and Biomedical engineering, 2009: p. 503-506.
14. Bontempi, M., et al., *Monte Carlo simulator of realistic x-ray beam for diagnostic applications*. Medical Physics, 2010. **37**(8): p. 4201-4209.
15. Poludniowski, G., et al., *SpekCalc: a program to calculate photon spectra from tungsten anode x-ray tubes*. Phys Med Biol, 2009. **54**(19): p. N433-8.

16. Poludniowski, G.G., *Calculation of x-ray spectra emerging from an x-ray tube. Part II. X-ray production and filtration in x-ray targets.* Med Phys, 2007. **34**(6): p. 2175-86.
17. Poludniowski, G.G. and P.M. Evans, *Calculation of x-ray spectra emerging from an x-ray tube. Part I. electron penetration characteristics in x-ray targets.* Med Phys, 2007. **34**(6): p. 2164-74.
18. Punnoose, J., et al., *Technical Note: spektr 3.0-A computational tool for x-ray spectrum modeling and analysis.* Medical Physics, 2016. **43**(8Part1): p. 4711-4717.
19. Agostinelli, S., et al., *Geant4—a simulation toolkit.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2003. **506**(3): p. 250-303.
20. Rong, X.J., et al., *Measurement of focal spot size with slit camera using computed radiography and flat-panel based digital detectors.* Med Phys, 2003. **30**(7): p. 1768-75.
21. Baer, R.L., *The circular-edge spatial frequency response test.* Proceedings of SPIE - The International Society for Optical Engineering, 2003.
22. Boone, J.M., T. Yu, and A. Seibert, *Mammography spectrum measurement using an x-ray diffraction device.* Phys Med Biol, 1998. **43**(9): p. 2569-82.
23. Kimme-Smith, C., et al., *Mammograms obtained with rhodium vs molybdenum anodes: contrast and dose differences.* AJR Am J Roentgenol, 1994. **162**(6): p. 1313-7.
24. Granton, P.V. and F. Verhaegen, *On the use of an analytic source model for dose calculations in precision image-guided small animal radiotherapy.* Phys Med Biol, 2013. **58**(10): p. 3377-95.
25. van Hoof, S.J., J.B. Verde, and F. Verhaegen, *Dose painting by dynamic irradiation delivery on an image-guided small animal radiotherapy platform.* Br J Radiol, 2019. **92**(1095): p. 20180744.
26. Alaei, P. and E. Spezi, *Imaging dose from cone beam computed tomography in radiation therapy.* Phys Med, 2015. **31**(7): p. 647-58.
27. Ding, G.X., et al., *Image guidance doses delivered during radiotherapy: Quantification, management, and reduction: Report of the AAPM Therapy Physics Committee Task Group 180.* Med Phys, 2018. **45**(5): p. e84-e99.
28. Létang, J.-M. and G. Peix, *On-line X-ray focal spot assessment based on deconvolution using standard imaging devices.* NDT & E International, 2003. **36**(5): p. 303-317.
29. Nagesh, S.V.S., et al., *Focal Spot Deblurring for High Resolution Direct Conversion X-ray Detectors.* Proc SPIE Int Soc Opt Eng, 2016. **9783**.
30. Kuntz, J., et al., *Focal spot deconvolution using convolutional neural networks.* Proc SPIE Medical Imaging 2019: Physics of Medical Imaging, 2019. **10948**.

Supplementary Materials

A. Backscatter, entry, and X-ray production ratios for a 20-degree anode angle

Table S1. Small focal spot backscatter, entry, and X-ray production ratios for different generations and three X-ray tube potentials (80, 100, and 120 kV) for a 20-degree anode angle.

Backscatter ratio (% of the 1 st entry)	80 kV		100 kV		120 kV	
	<i>Small</i>	<i>Large</i>	<i>Small</i>	<i>Large</i>	<i>Small</i>	<i>Large</i>
1 st generation	49.4		49.8		49.9	
2 nd generation	5.5	5.3	5.7	5.6	5.9	5.8
Entry ratio (% of the 1st entry)						
2 nd entry	8.1	7.9	8.4	8.2	8.7	8.4
3 rd entry	1.3	1.2	1.4	1.4	1.5	1.5
X-ray production ratio (% of the 1st entry)						
1 st generation	1.10		1.65		2.29	
2 nd generation	0.02		0.03		0.05	

7

Chapter 7

A Monte Carlo based scatter removal method for non-isocentric cone-beam CT acquisitions using a deep convolutional autoencoder

Brent van der Heyden*, Martin Uray*, Gabriel P Fonseca, Philipp Huber, Defne Us, Ivan Messner, Adam Law, Anastasiia Parii, Niklas Reisz, Ilaria Rinaldi, Gloria Vilches-Freixas, Heinz Deutschmann, Frank Verhaegen, and Philipp Steininger

* Equal contribution

Submitted work

Abstract

Background and Purpose: In cone-beam computed tomography (CBCT) imaging, the primary imaging beam scatters inside the patient and produces a contaminating photon fluence that is registered by the detector. Scattered photons cause artifacts in the image reconstruction, and are partially responsible for the inferior image quality compared to diagnostic fan-beam CT. In this work, a deep convolutional autoencoder (DCAE) and projection-based scatter removal algorithm were constructed for the ImagingRing™ system on rails (IRr), which allows for non-isocentric acquisitions around virtual rotation centers with its independently rotatable source and detector arms.

Materials and Methods: An IRr Monte Carlo model (Geant4) was developed and used to simulate a training dataset of approximately 1280 projection pairs (primary + scatter) from 27 digital head-and-neck cancer patients around five different virtual rotation centers. The scatter removal performance was investigated in two digital anthropomorphic phantom simulations and in eight real patient acquisitions. Measures for the quantitative error, the signal-to-noise ratio, and the similarity were evaluated for two simulated digital head-and-neck patients, and the contrast-to-noise ratio (CNR) was investigated between muscle and adipose tissue in the real patient image reconstructions. Image quality metrics were compared between the uncorrected data, the currently implemented heuristic scatter correction data, and the DCAE corrected image reconstruction.

Results: The DCAE corrected image reconstructions of two digital patient simulations showed superior image quality metrics compared to the uncorrected and corrected image reconstructions using a scatter removal heuristic. The proposed DCAE scatter correction in this study was successfully demonstrated in real non-isocentric patient CBCT acquisitions and achieved statistically significant higher CNRs compared to the uncorrected or the heuristic corrected image data.

Conclusion: This thesis chapter presents a projection-based scatter removal algorithm for CBCT imaging using a deep convolutional autoencoder trained on Monte Carlo composed datasets. The algorithm was successfully applied to real patient data.

Conflict of Interest

The authors disclosed receipt of joint financial support (Eurostars - MEIRS, reference number 11605; FFG, reference 861339; RVO, Rijksdienst voor Ondernemend Nederland) for the research of this thesis chapter.

Acknowledgements

The authors thank Dr. Richard Canters, medical physicist at Maastrou, for his valuable support during this study.

Introduction

Cone-beam computed tomography (CBCT) is frequently used in image-guided photon radiotherapy [1] and recently found its way to proton radiotherapy as alternative for diagnostic in-room CT [2, 3]. The clinical potential of CBCT in photon radiotherapy has already been demonstrated for patient positioning [4], adaptive treatment planning [5], and dose recalculation techniques [6]. However, the CBCT image quality is still inferior to the diagnostic planning CT image quality, which is conventionally used in radiotherapy treatment planning, for e.g. the organ-at-risk segmentation and the dose calculations.

The inferior image quality of CBCT systems is a consequence of different limitations such as; the coherent and the incoherent photon scatter contributions [7], flat-panel detector readout artifacts such as image lag [8-10], and patient motion [11, 12]. Of these, patient motion and scattered photon fluences are two major aspects that significantly influence the CBCT image quality [13].

Photon and proton radiotherapy workflows would both benefit from improved CBCT image quality and therefore several software and hardware solutions have been proposed to reduce CBCT motion artifacts by sinogram interpolation techniques or 4D image reconstruction techniques [14-17], and to correct for scattered particle distributions at the level of the detector [13, 18-22]. These corrections are desired to improve image contrast and Hounsfield Unit (HU) consistency while reducing image noise and image reconstruction artifacts (e.g. motion or beam hardening).

The interest in artificial intelligence algorithms for radiotherapy applications is growing rapidly due to its many use cases, and recently these novel algorithms have been proposed as scatter removal solutions in projection and image space. Maier *et al.* [23] demonstrated a real-time deep scatter estimation method which was trained to reproduce the Monte Carlo scatter simulation output based on acquired projection data of mechanical system components (e.g. a compressor wheel). Kida *et al.* [24] proposed an image-based deep learning approach that was trained on prostate patient data to perform direct image intensity mapping from CBCT to the registered planning CT image quality. Hansen *et al.* [25] presented a hybrid method that was trained to remove scatter distributions based on forward projected data of planning CTs that were deformably registered to the acquired CBCT.

In Landry *et al.* [26], three different U-Net algorithms, including previously published works [24, 25], were compared to adjust CBCT images for prostate radiotherapy dose calculations treated by proton pencil beam scanning and by photon volumetric arc therapy. Nomura *et al.* [27] also constructed a U-Net to correct for CBCT X-ray scatter. Their training dataset consisted of 1800 X-ray projections acquired with fixed source-to-detector distance from a limited number of non-anthropomor-

phic phantoms and their correction algorithm was validated on anthropomorphic phantom data and digital lung phantoms. More recently, Lee and Lee [28] adopted a convolutional neural network approach to restore CBCT projections based on Monte Carlo simulations.

These approaches to correct for photon scatter in projection and image space are only evaluated on isocentric projection data. In the literature, approaches to correct for scatter in non-isocentric CBCT trajectories with varying source-to-detector distances are currently lacking. Non-isocentric CBCT imaging can be achieved by imaging modalities with an independently rotating source and detector to focus on geometries of which the geometrical rotation center is not aligned with the mechanical isocenter of the imaging modality. In either isocentric or non-isocentric imaging protocols, small and large field-of-view trajectories can be adopted for imaging of smaller body sites (e.g. neuro) and larger body sites (e.g. lung and breast). Small field-of-view (SFOV) trajectories cover 360-degrees with the detector panel centered on the virtual rotation center of the imaging trajectories, large field-of-view (LFOV) trajectories also cover 360-degrees, but with a shifted detector position to cover a larger overall volume.

In this work, a full Monte Carlo simulation model of the X-ray tube and imaging panel was developed to simulate primary and scattered photon transmissions through digital patient CT geometries. Non-isocentric simulated X-ray projection pairs (i.e. primary transmission and scattered transmission) were used to train a deep scatter removal algorithm. The algorithm performance was first tested and validated on simulated Monte Carlo transmission values through realistic digital patient geometries based on fan-beam CT data and then on real patient non-isocentric projection data.

Materials and methods

Scatter correction workflow

In the scatter correction workflow (Figure 1), the total photon distribution measured by the flat panel detector was described as the sum of the primary photon distribution (p_{prm}) and the scattered secondary photon distribution (p_{sc}). In non-isocentric CBCT trajectories of the ImagingRingTM (medPhoton GmbH, Salzburg, Austria), the simulated and the experimental X-ray projections $p(S, D)$ depend on both the source position S and the detector position D . The input projections, either simulated or measured experimentally, were first downsampled from a 1024x1024 grid to a 256x256 image grid, and the negative logarithm for each pixel was calculated subsequently to obtain suited input images for 3D reconstruction. This downsampled input projections were then used as input for the trained scatter prediction model to

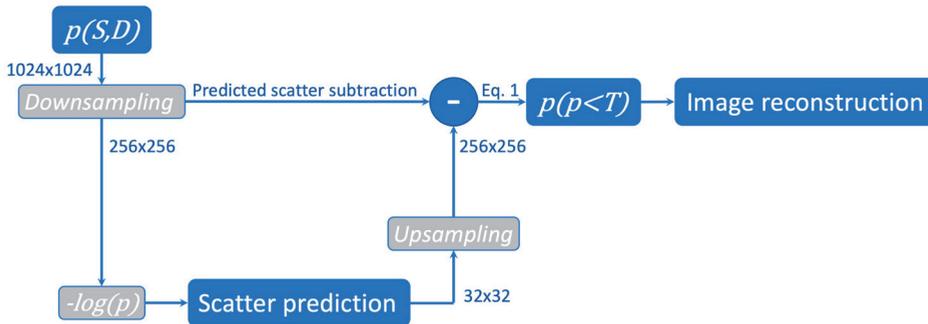


Figure 1. Flowchart of the projection-based scatter correction workflow starting with downsampling the measured projection acquired with source position S and with detector position D . Then, the scatter is predicted and subtracted from the downsampled projection (256×256). Finally, post processing is applied on the corrected projection (Equation 1), whereafter the CBCT image reconstruction is performed.

predict a 32×32 scatter distribution. In preceding experiments, the downsampling and upsampling of the scatter distributions showed no information loss from the original signal (Supplementary Material A). Predicted scatter distributions were bilinearly upsampled to 256×256 to meet the downsampled input projection size. Next, the predicted and upsampled scatter distribution data was subtracted from the input projection data to obtain scatter-free X-ray projections. After scatter subtraction, image post processing was required in projection space (Equation 1) to correct for small negative projection values resulting from a scatter overestimation in high transmission areas. In Equation 1, an empirically determined constant threshold was employed ($T=0.015$), which provided the best correction results on projection level (Supplementary Material B). Finally, the corrected projection data was reconstructed using the Feldkamp-Davis-Kress (FDK) filtered back-projection algorithm.

$$p(p < T) = T \cdot e^{\frac{p-T}{T}} \quad \text{where } T = 0.015 \quad (\text{Eq. 1})$$

Network architecture

A neural network architecture was built using Keras v.2.2 [29] (Tensorflow 1.13 and CUDA 10) using a deep convolutional autoencoder (DCAE). The architecture was composed of an encoder and a decoder (Figure 2). Each layer in the encoder component consists of three two-dimensional convolutional layers directly followed by a Rectified Linear Unit (ReLU) and a max pooling layer. A layer in the decoder component consists of two convolutional layers directly followed by a bilinear upsampling layer, which doubled the size in each dimension. Image padding was applied in all convolutional layers and the kernel size was set to 3×3 . The number of filters was set to $n_i = [4, 40, 80, 160, 320, 480, 960]$, where i denotes the depth of the network archi-

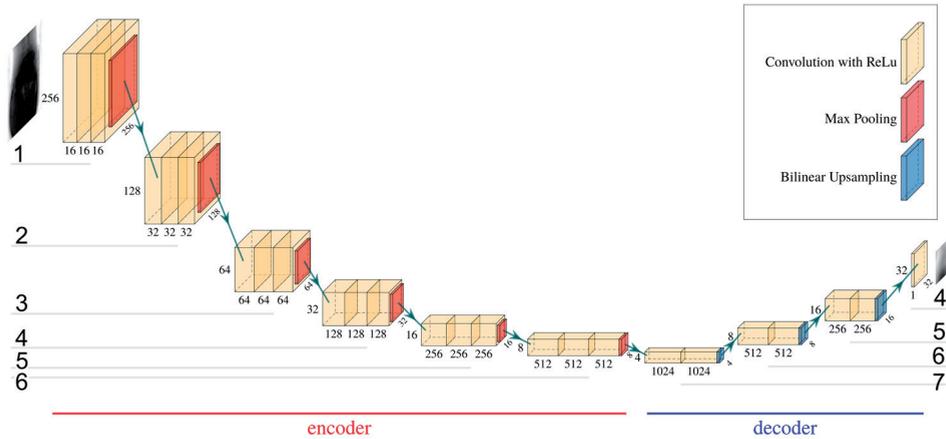


Figure 2. Architecture diagram of the deep convolutional autoencoder (DCAE). The horizontal grey lines indicate the depth of the network architecture i (1-7). (Abbreviation: ReLu, rectified linear unit).

ture which is indicated by the horizontal grey lines in Figure 2. Compared to the U-net inspired network architectures used in previous works such as Landry *et al.* [26] or Nomura *et al.* [27], this deep convolutional autoencoder does not employ skip-connections. These links connect the encoder and the decoder at the same level, by concatenating the representation at the end of the encoding step to the output of the decoding step one level below. This fosters the training and prediction of high-frequency features, as the high frequency information from the input is directly passed to the last layer.

For training, Adam was employed as optimizer (parametrized with $\beta_1 = 0.900$ and $\beta_2 = 0.999$) with an initial learning of $\alpha = 1e - 5$ [30]. A weight decay was configured to halve the learning rate after 40 epochs. The batch size was set to 12. Data augmentation was applied according to the *mixup* technique [31] to facilitate generalization of the network. In this *mixup* technique, the weighted sum of two independent samples (projections) was used to generate additional, virtual augmented data for training. The weighting factor was randomly sampled from a Beta distribution ($\alpha = 0.2, \beta = 0.2$). Parameter tuning of the architecture was investigated empirically to reduce the risk of overfitting. Training was performed on a computer equipped with a NVIDIA GeForce RTX 2080 Ti GPU and an Intel Core i9-9820X CPU.

Training dataset and Monte Carlo simulation data

A Monte Carlo simulation model of the ImagingRing™ system on rails (IRr) was built in Geant4 (v10.5) [32]. The simulation model consisted of a finite polychromatic X-ray source, a patient geometry, and a flat-panel detector (Figure 3). The IRr system

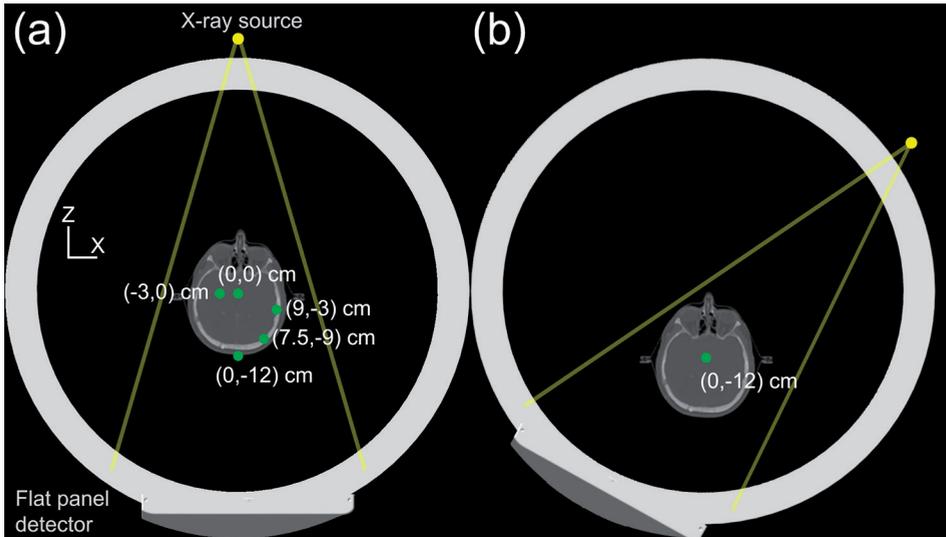


Figure 3. In the left panel (a), a sketch shows the Monte Carlo simulation geometry of the IRr scanner. The green dots indicate the virtual rotation centers of the simulated acquisition trajectories. Here, (0,0) cm is the volumetric isocenter of the digital patient phantom and thus the virtual rotation center of the CBCT imaging protocol. In the right panel (b), (0,-12) cm is the virtual rotation center of a non-isocentric imaging trajectory.

is a novel CBCT device that is capable of performing non-isocentric image acquisitions (Figure 3b) due to its independently rotatable X-ray source and flat panel-detector. System specifications required for the development of the Monte Carlo model were provided by the IRr system manufacturer.

SpekCalc [33-35] was used to calculate the 120 kV X-ray spectrum of the IRr system based on system specifications such as the anode angle (10 degrees) and the inherent tube filtration (1.3 mm Al). The IRr system has a built-in hardware flattening filter with varying material thickness (0-6 mm Al) to compensate for the heel effect, and therefore no heel effect simulations were performed. Technical drawings of the XRD 1642 Perkin Elmer flat-panel detector (40.96 x 40.96 cm², 1024 x 1024 px²) with CsI scintillator were used to calculate the detector energy response curve of the detector. Geant4's G4UserSteppingAction and G4VUserTrackInformation classes categorized X-rays as primary or scattered photons so that the imaging panel could discriminate between primary (p_{prm}) and scattered photon (p_{sec}) distributions, which were then used for training the scatter removal algorithm. Details about this detector simulation model were provided as [Supplementary Material C](#).

Non-isocentric and SFOV IRr imaging protocols with detailed source and detector positions were used in our model to rotate the source and the detector independently from each other. The training dataset consisted of 27 diagnostic head-and-

neck CT scans and brain CT scans obtained from freely available online resources (MICCAI auto segmentation challenge 2015 [36]). The head-and-neck treatment site was selected in this study because maximum CBCT image quality enhancements are desired to improve the accuracy of proton dose recalculations. These CT scans were used as ground-truth patient anatomy input of our IRr simulation model. Each patient was simulated with a SFOV non-isocentric or isocentric CBCT imaging protocol of 31 or 61 angularly equally spaced projections per patient, assigned randomly, and with respect to the virtual rotation center. Several IRr imaging protocols were simulated around five different virtual rotation centers (X, Z denoting a coordinate in the trans-axial plane): (0.0, 0.0) cm, (0.0, -12.0) cm, (-3.0, 0.0) cm, (9.0, -3.0) cm, and (7.5, -9.0) cm. The volume centers of the digital patient phantoms were aligned around the virtual rotation centers of the simulated imaging protocols and all patients were virtually positioned head-first-supine (Figure 3). Small intentional shifts of the volumetric digital patient isocenter and the virtual rotation isocenter were performed to simultaneously simulate truncated projection data, as acquired with LFOV imaging protocols. In total, 1287 simulated projection pairs (primary and scatter) were considered as training dataset for the network architecture. These projections were acquired with 5 different imaging protocols, and thus with 5 different virtual rotation centers, of 27 different digital head-and-neck patients.

Test dataset

The performance of the network architecture was first tested on simulated patient projections and then on real patient projections. For the simulated data, the CBCT image reconstruction of the primary transmission was compared against the uncorrected reconstruction of the combined transmission (primary + scatter), the reconstruction of the combined transmission after heuristic scatter correction [37], and the reconstruction of the combined transmission corrected by our DCAE network architecture. The heuristic approach described in [Supplementary Material D](#) is implemented in the commercial IRr software and corrects for photon scatter in X-ray projections that are acquired with non-isocentric CBCT protocols.

Simulation data for model validation

The same Monte Carlo model that was used to generate the training dataset, was likewise used to create simulation data for model testing. Two diagnostic brain CT images, not included in the training dataset, were converted to digital Monte Carlo patient geometries. The virtual rotation centers of the IRr imaging protocols were (0.0, -12.0) cm and (-3.0, 0.0) cm. To evaluate the performance of the scatter removal algorithm in image space, a sufficient number of projections was required for the CBCT image FDK-based reconstruction to prevent artifacts induced by angularly sparsely sampled data. Therefore, 361 projections were simulated per digital patient geometry.

Retrospective patient data

The retrospective study followed the regulatory guidelines and was approved by the internal review board of Maastrro (Maastricht, the Netherlands). Eight head-and-neck cancer patients were treated with proton radiotherapy at Maastrro Protonentherapie (Maastricht, the Netherlands), and were randomly included in this study. Before the image guided proton treatment fraction, LFOV imaging with the IRr system was performed to image the patient, including shoulders, for positioning. The logs and the raw projection data were exported offline to test the scatter removal algorithm independent from the clinical workflow.

Quantitative evaluation metrics

The simulated X-ray projections (n=361) of the non-isocentric imaging protocols were reconstructed for two digital patient phantoms. Five quantitative metrics, including the contrast-to-noise ratio (CNR), the mean absolute error (MAE), the mean square error (MSE), the peak signal-to-noise ratio (PSNR), and the structural similarity (SSIM) index [38] were adopted to evaluate the reconstructed CBCT quality in image space. The CNR ratio between the mean muscle \bar{I}_{muscle} and the mean adipose \bar{I}_{adipose} image intensity was calculated in the same region-of-interest for every CBCT image reconstruction, which was manually delineated in the neck of every patient (Equation 2). The σ_{muscle} and σ_{adipose} are the standard deviations of \bar{I}_{muscle} and \bar{I}_{adipose} . Statistical significance between the CNRs calculated for the different CBCT reconstructions of eight patients in the retrospective study was evaluated with a paired t-test.

$$\text{CNR} = \frac{\bar{I}_{\text{muscle}} - \bar{I}_{\text{adipose}}}{\sqrt{\sigma_{\text{muscle}}^2 + \sigma_{\text{adipose}}^2}} \quad (\text{Eq. 2})$$

Four other metrics were calculated relative to the primary CBCT reconstruction, which was considered as reference for the simulated digital patient phantoms in this study. These four metrics were evaluated three-dimensionally over all slices (excluding air) between the primary reconstruction as reference, and the uncorrected, the heuristic-corrected, and the DCAE corrected CBCT image reconstructions. The MAE, MSE, PSNR, and the SSIM index were calculated by Equations 3-6.

$$\text{MSE} = \frac{1}{N} \sum_n (I_{\text{prm}} - I_X)^2 \quad (\text{Eq. 4})$$

$$\text{PSNR} = 10 \log_{10} \frac{\max(I_X)^2}{\text{MSE}} \quad (\text{Eq. 5})$$

$$\text{SSIM} = \frac{(2 \mu_{\text{prm}} \mu_X + c_1)(2 \sigma_{\text{prm},X} + c_2)}{(\mu_{\text{prm}}^2 + \mu_X^2 + c_1)(\sigma_{\text{prm}}^2 + \sigma_X^2 + c_2)} \quad (\text{Eq. 6})$$

where, $c_1 = (0.01/1E9 \text{ l})^2$ and $c_2 = (0.03/1E9 \text{ l})^2$

In the equations, I_{prm} denotes the vectorized primary image reconstruction (in 1/cm), I_X is the vectorized image of reconstruction method X, N is the total number of voxels, μ is the mean of the image vector, $\sigma_{\text{prm},X}$ is the covariance of the primary image vector and the image vector of method X, and σ^2 is the variance of the image vector. Here, l is the dynamic range of the pixel values ($l = 2^{32} - 1$, for 32-bit floats). The 32-bit image values in the vector images are linear attenuation coefficients and are relatively small (<1), hence the default SSIM constants 0.01 and 0.03 are both divided by 1E9 (Equation 6).

Results

Training and Monte Carlo dataset

During the training phase of the network architecture, the MAE metric was used as loss function (Figure 4). The computation time required to train the network on the simulated Monte Carlo X-ray projections (examples in Figure 5) was 464 seconds, and the average prediction time for one projection was 4.35 milliseconds on the GeForce RTX 2080 Ti GPU. The heuristic scatter correction approach required on average 4.04 milliseconds per projection.

The primary and scattered photon distributions in one projection simulated for three example patients, randomly selected out of the training dataset, are shown in Figure 5. The source-detector position (in degrees) and the virtual rotation center are reported for these patients.

Two digital patient CBCT reconstructions

The diagnostic CT and the primary, uncorrected, heuristic corrected, and the DCAE corrected CBCT image reconstructions for the first head-and-neck patient CBCT simulation with 361 X-ray projections are shown in Figure 6. Similar to Figure 6, Figure 7 depicts the CBCT image reconstructions and cross-profiles for the second digital head-and-neck cancer patient at a different longitudinal slice position. The Elastix deformable image registration toolbox [39] was only used for visualization purposes to register the diagnostic fan-beam CT with the CBCT image reconstructions.

The quantitative metrics to compare the image quality against the reference (primary reconstruction) were calculated for both digital head-and-neck cancer patient phantoms and are listed in Table 1. These metrics are calculated over all slices, excluding air voxels from the analysis with simple thresholding: $\mu < 0.02 \text{ cm}^{-1}$.

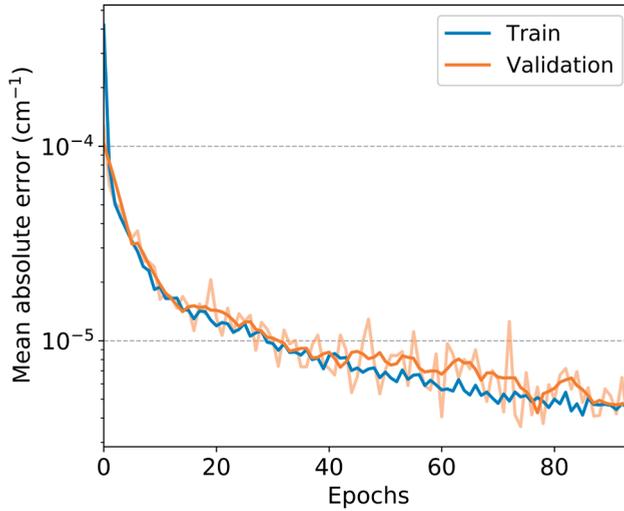


Figure 4. Mean absolute error (MAE) loss function plotted against the number of epochs for the raw training data (blue) and the smoothed validation data (orange). The raw MAE values of the validation loss curve are plotted with a higher transparency.

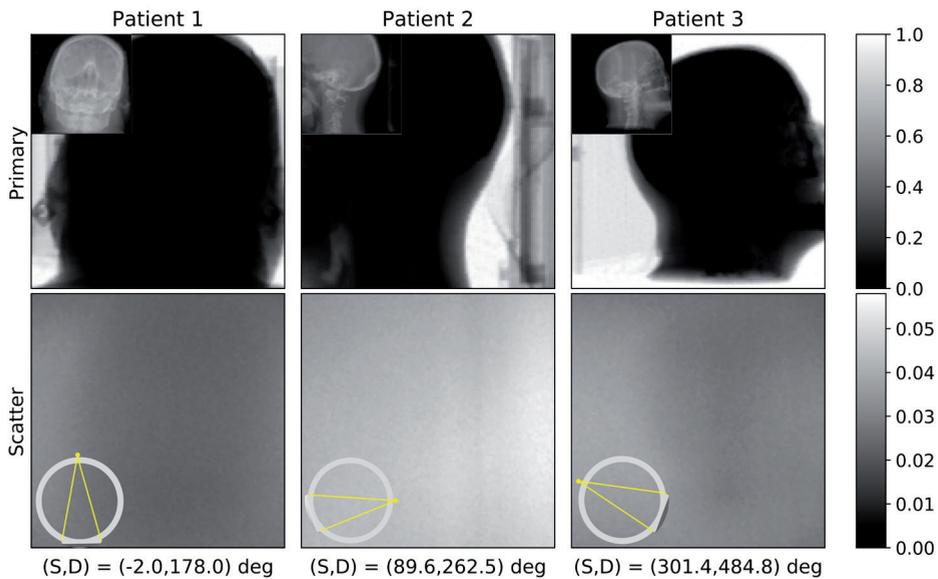


Figure 5. X-ray projection pairs in transmission units for three example patients simulated by Monte Carlo. The source-detector positions (S, D) in degrees were reported for each of the projections. Digital example patient 1 was imaged around virtual rotation center (0.0, 0.0) cm, patient 2 around (9.0, -3.0) cm, and patient 3 around (7.5, -9.0) cm.

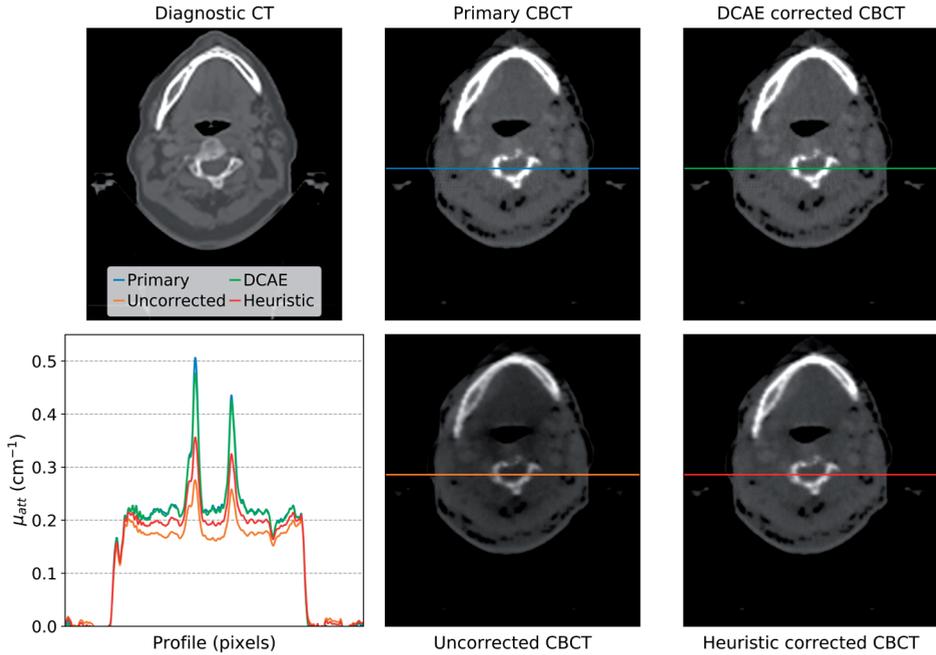


Figure 6. The image intensity cross-profiles of the linear attenuation coefficient through the primary (blue), the uncorrected (orange), the heuristic corrected (red), and the deep convolutional autoencoder (DCAE) corrected (green) CBCT image reconstructions for the first digital head-and-neck patient simulated around virtual rotation center (0, -12) cm. The primary and DCAE cross-profiles overlap over the majority of the profile. (window width/level =0.17/0.27 cm⁻¹), whereas significant differences can be seen for the other methods.

Table 1. Quantitative image quality comparison for two digital head-and-neck cancer patient reconstructions with the mean absolute error (MAE), the mean square error (MSE), the peak signal-to-noise ratio (PSNR), the structural similarity index (SSIM), and the contrast-to-noise ratio (CNR) between muscle and adipose tissue.

		Uncorrected	Heuristic Corrected	DCAE corrected	Primary
Digital Patient 1	MAE [cm ⁻¹]	57.7E-3	32.2E-3	5.7E-3	-
	MSE [cm ⁻²]	5.5E-3	1.9E-3	0.1E-3	-
	PSNR [dB]	17.1	22.5	36.5	-
	SSIM [-]	0.896	0.962	0.997	-
	CNR [-]	0.73	1.41	2.34	2.81
Digital Patient 2	MAE [cm ⁻¹]	59.5E-3	48.8E-3	5.9E-3	-
	MSE [cm ⁻²]	5.9E-3	3.9E-3	0.1E-3	-
	PSNR [dB]	17.1	19.3	37.1	-
	SSIM [-]	0.896	0.933	0.997	-
	CNR [-]	1.83	1.87	2.03	2.34

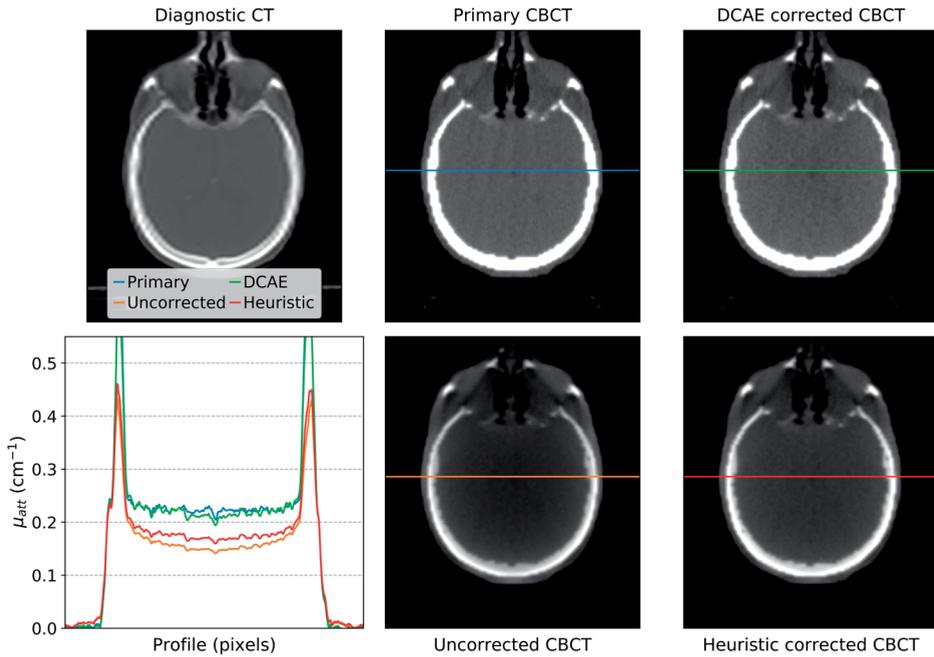


Figure 7. The image intensity cross profiles of the linear attenuation coefficient through the primary (blue), the uncorrected (orange), the heuristic corrected (red), and the deep convolutional autoencoder (DCAE) corrected (green) CBCT image reconstructions for the second digital head-and-neck patient simulated around virtual rotation center $(-3,0)$ cm. (window width/level = $0.17/0.27$ cm^{-1})

Retrospective patient study

The non-isocentric CBCT image reconstructions of the uncorrected, the heuristic corrected, and the DCAE corrected X-ray projections acquired in real patient acquisitions are depicted in [Figure 8](#). All reconstructed CBCT images were converted to Hounsfield Units to compare with the diagnostic radiotherapy planning CT. All diagnostic CTs (in Hounsfield Units) were registered deformably with the CBCT reconstruction to evaluate cross-profile differences [39]. The CNRs between muscle and adipose tissues, calculated according to [Equation 2](#), are listed in [Table 2](#). Calculated over the entire retrospective patient cohort, the CNRs achieved with the DCAE correction algorithm were significantly higher than the CNRs calculated for the heuristic corrected ($p=0.017$) and the uncorrected ($p=0.014$) CBCT reconstructions.

Table 2. The contrast-to-noise ratio (CNR) calculated for the planning CT and for the uncorrected, the heuristic corrected, and the DCAE corrected non-isocentric CBCT image reconstructions of eight patient acquisitions.

CNR [-]	Uncorrected	Heuristic Corrected	DCAE corrected	Planning CT
Patient 1	1.6	2.5	2.5	7.7
Patient 2	1.9	3.3	5.0	6.9
Patient 3	2.6	2.9	3.0	8.0
Patient 4	2.9	2.5	3.1	3.4
Patient 5	3.6	4.9	5.0	5.8
Patient 6	2.5	2.6	2.9	3.0
Patient 7	3.3	3.5	3.5	4.1
Patient 8	2.0	2.0	2.6	3.4

Discussion

X-rays that scatter inside the patient contribute in a major way to the reduced CBCT image quality compared to diagnostic CT [8]. Algorithms or hardware to correct for the scattered photon fluence are desired to improve CBCT image quality. This study describes the development and the validation of a projection-based scatter removal algorithm for CBCT using a deep convolutional autoencoder. For network training, a projection dataset of primary and scattered photon transmissions was provided based on Monte Carlo simulations of realistic digital head-and-neck cancer patient geometries. The algorithm’s performance was first evaluated on simulated data and then on real acquired patient data. The head-and-neck body site was chosen to challenge the performance of the DCAE scatter correction algorithm, due to the heavy varying body thickness (e.g. neck and shoulders), and therefore, variations in object scatter.

The non-isocentric CBCT image reconstruction of two simulated digital head-and-neck patients showed superior performance of the DCAE scatter correction algorithm compared to the uncorrected and heuristic corrected image reconstruction relative to the image reconstruction of the primary photon transmission. Error metrics such as the MAE and MSE reported the lowest errors for the DCAE corrected image reconstruction (Table 1). The MAEs calculated for the DCAE corrected image reconstructions (0.0057 cm^{-1} and 0.0059 cm^{-1}) were even a magnitude lower than the uncorrected image reconstructions (0.0577 cm^{-1} and 0.0595 cm^{-1}). Also, the MSEs calculated for the DCAE corrected image reconstructions were approximately 50 times lower compared to the uncorrected and approximately 29 times lower compared to the heuristic corrected image reconstructions. The PSNR and SSIM metrics confirmed that the DCAE corrected image reconstruction achieved better CBCT image quality.

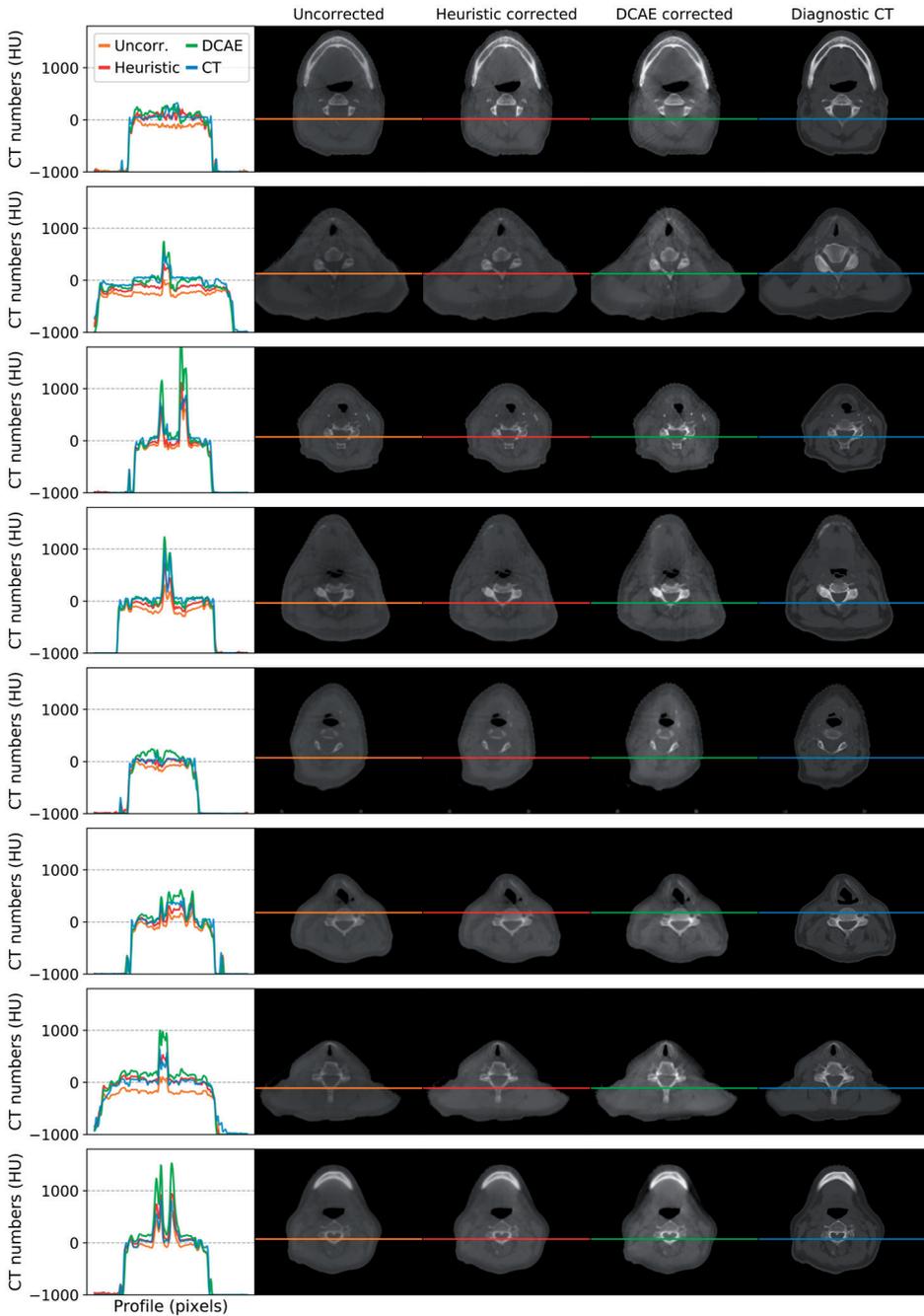


Figure 8. Axial CBCT image reconstruction for the uncorrected (orange), heuristic corrected (red), and the DCAE corrected (green) CBCT image reconstructions in eight real patient acquisitions. The right most column shows the deformably registered diagnostic CTs (blue) acquired prior to proton treatment planning.

After testing the model on simulated digital patient data, the trained DCAE model was applied on non-isocentric CBCT patient acquisitions acquired with LFOV IRr imaging protocols, which challenges the DCAE model even more. Here, no primary photon transmission distributions are available as ground truth and, therefore, the CNR was calculated between manually segmented muscle and adipose tissue regions in the neck of every patient. The DCAE corrected CBCT image reconstructions showed the highest CNR for all patients and was superior compared to the uncorrected and the heuristic corrected image reconstructions which showed lower CNRs. The highest CNRs were calculated for the planning CT due to the lower HU standard deviations in the of muscle and adipose regions-of-interest. A subjective image quality comparison between the simulated and the acquired CBCT image reconstructions showed more visible image artifacts (e.g. streak artifacts) in the actual patient acquisitions because these are subject to patient motion and beam hardening artifacts. Streak artifacts were observed in all patient image reconstructions; however, these streaks were more highlighted in the DCAE corrected compared to the uncorrected and heuristic corrected image reconstructions due to improved image contrast (CNR).

In clinical routine, CBCT based patient alignment is often performed prior to radiotherapy treatment and the image reconstruction times are of interest because the CBCT is reconstructed online. On average, the IRr imaging protocols (factory setting) acquire 600 X-ray projections per acquisition. Here, the heuristic scatter correction algorithm required 4.04 milliseconds per projection or approximately 2.4 seconds per image reconstruction. The DCAE scatter correction method presented in this study was 8% slower and requires 4.35 milliseconds per projection or approximately 2.6 seconds per image reconstruction. Inline CBCT image reconstruction and acquisition times (≈ 1 minute) required much longer than the scatter correction algorithm, which makes the correction speed of both scatter removal algorithms comparable in a clinical radiotherapy setting.

In recent years, a growing body of literature investigated the use of deep learning to remove photon scatter from CBCT projections. Currently, these algorithms are either trained on Monte Carlo simulation data or on forward projected data from deformed diagnostic CT scans. Nomura *et al.* [27] and Lee and Lee [28] both presented a network that was trained and validated on Monte Carlo simulation data. However, neither of these Monte Carlo based studies tested their network architecture on real acquired patient data. Other studies investigated the use of diagnostic CT scans as ‘scatter-free’ ground truth to generate a training dataset [25] or trained a neural network to map the CBCT image to the deformed diagnostic CT image quality in image space [24]. In contrast to the Monte Carlo based approaches, Kida *et al.* [24] and Hansen *et al.* [25] applied their methods on real patient data. Earlier works [23, 27] used neural network architectures (U-Net) that preserved high

frequencies from the network input to the prediction output which is one of the requirements for automatic biomedical image segmentation tasks [40]. During the preliminary experiments of this work, originally performed with a U-Net architecture, structure boundaries were observed in the predicted scatter output that resembled the body outline of the input projection. It is likely that earlier works used more projection training data, but still a lower number of patients, to overcome these outline artifacts and to better account for low frequencies in the predicted projection output. This work finally adopted a reduced network architecture complexity to facilitate the identification of the dominant low-frequency scatter characteristics without the presence of body outlines in the predicted scatter output, which simultaneously enabled faster training (≈ 8 minutes) with a minor demand of Monte Carlo training data.

Earlier studies already developed Monte Carlo based scatter correction algorithms, however, the present study shows novelty because, to our best knowledge, this is the first Monte Carlo based scatter removal algorithm in CBCT imaging that is applied on real patient acquisitions. Furthermore, our method also applies to non-isocentric imaging protocols with varying source-to-detector distances in one imaging protocol. These varying source-to-detector distances directly affect the scatter distributions and are handled by the algorithm without prior knowledge (e.g. header information). In future efforts to improve the performance of the DSAE algorithm, additional channels in the network could be provided with the source-to-detector and detector-to-object distance based on header information. However, our simulated training dataset only consisted of head-and-neck cancer patients, which makes it difficult to estimate how well the scatter removal method will perform on other body sites that generate more scatter such as the pelvis or the thorax. For these larger body sites, the method accuracy needs to be investigated separately.

The prevalence of CBCT imaging in radiotherapy gave rise to image-guided radiotherapy and adaptive radiotherapy. Methods to improve CBCT quality have the potential to improve the photon and proton radiotherapy workflow, especially in proton therapy, where better CBCT image quality could lead to the clinical implementation of online dose recalculations [41, 42]. Most of the applications benefit from the scatter removal approach after the CBCT reconstruction in image space, but scatter corrected projections are also promising for projection-based methods, e.g. the stopping power estimations with dual-energy CT [43].

The imaging protocols in this study were all simulated with a conventional 120 kV X-ray spectrum which can introduce a bias in this study. It needs to be further investigated how this algorithm will perform when changes are applied to the imaging protocols: e.g. X-ray tube potential and spectral filtration. Further research is required to investigate the model robustness and provide a solution that is applicable to a broader range of imaging specifications. It is likely that the diversity in the

training dataset needs to be increased or that multiple protocol-specific models need to be trained. Projection header information provides more information about the imaging protocol and thus could be used to determine which specific model to apply. Incorporating header information from the individual projections (e.g. kV, filtration, object-to-detector distance) in future studies would allow to integrate fast-kV switching and dual-energy CBCT correction algorithms to improve projection-based parameter estimations.

In radiotherapy, no affordable detector technology is currently available to detect and separate primary and scattered photons in CBCT projections, which can be used as training dataset for deep learning network architectures. Therefore, CBCT Monte Carlo simulations of digital patient phantoms provided a good solution to compose a training dataset, which can also be seen as a limitation of the presented scatter removal method. To perform Monte Carlo simulations, assumptions have to be made about the patient's tissue composition and density. The resolution of the Monte Carlo simulation geometry is also limited because of the assumed digital patient voxelization is only an approximation to the real patient geometry. Currently, software solutions aim to remove scatter at the projection level and hardware solutions (e.g. anti-scatter grids) aim to stop scattered particles before they are detected by the imaging panel, but in future applications, more affordable photon counter detectors such as MediPix3 [44, 45] or futuristic MediPix4 technology could correct for scattered particles as hybrid technology based on the measured particle incidence angle or the particle time-of flight.

Conclusion

This thesis chapter presented a novel projection-based scatter correction algorithm to remove the scattered photon transmission in CBCT projections using a deep convolutional autoencoder architecture (DCAE). Monte Carlo simulations of non-isocentric CBCT image acquisitions were performed to create a diverse training dataset of head-and-neck cancer patients. The DCAE method was successfully applied on simulated digital phantoms and real patient image acquisitions. Promising results were obtained for the non-isocentric DCAE corrected image reconstruction for head-and-neck cancer patients. All image reconstructions after DCAE scatter correction showed better contrast-to-noise ratios compared to the uncorrected and the heuristic corrected image reconstruction. Promising results were obtained, and to conclude, future research is desired to extend the simulated training dataset and perform model validation for more treatment sites to have a general working scatter correction model that has the potential to be implemented in the clinical workflow.

References

1. Verellen, D., et al., *Innovations in image-guided radiotherapy*. Nat Rev Cancer, 2007. 7(12): p. 949-60.
2. Hua, C., et al., *A robotic C-arm cone beam CT system for image-guided proton therapy: design and performance*. Br J Radiol, 2017. 90(1079): p. 20170266.
3. Stock, M., et al., *The technological basis for adaptive ion beam therapy at MedAustron: Status and outlook*. Z Med Phys, 2018. 28(3): p. 196-210.
4. Bell, K., et al., *Image guidance and positioning accuracy in clinical practice: influence of positioning errors and imaging dose on the real dose distribution for head and neck cancer treatment*. Radiat Oncol, 2018. 13(1): p. 190.
5. Hvid, C.A., et al., *Cone-beam computed tomography (CBCT) for adaptive image guided head and neck radiation therapy*. Acta Oncol, 2018. 57(4): p. 552-556.
6. Giacometti, V., et al., *An evaluation of techniques for dose calculation on cone beam computed tomography*. Br J Radiol, 2019. 92(1096): p. 20180383.
7. Siewerdsen, J.H. and D.A. Jaffray, *Cone-beam computed tomography with a flat-panel imager: effects of image lag*. Med Phys, 1999. 26(12): p. 2635-47.
8. Siewerdsen, J.H. and D.A. Jaffray, *Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter*. Med Phys, 2001. 28(2): p. 220-31.
9. Mail, N., P. O'Brien, and G. Pang, *Lag correction model and ghosting analysis for an indirect-conversion flat-panel imager*. J Appl Clin Med Phys, 2007. 8(3): p. 2483.
10. Poludniowski, G., et al., *Removal and effects of scatter-glare in cone-beam CT with an amorphous-silicon flat-panel detector*. Phys Med Biol, 2011. 56(6): p. 1837-51.
11. Zhang, Q., et al., *Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model*. Med Phys, 2010. 37(6): p. 2901-9.
12. Nardi, C., et al., *Motion artefacts in cone beam CT: an in vitro study about the effects on the images*. Br J Radiol, 2016. 89(1058): p. 20150687.
13. Ruhrnschopf, E.P. and K. Klingensbeck, *A general framework and review of scatter correction methods in x-ray cone-beam computerized tomography. Part 1: Scatter compensation approaches*. Med Phys, 2011. 38(7): p. 4296-311.
14. Marchant, T.E., et al., *Reduction of motion artefacts in on-board cone beam CT by warping of projection images*. Br J Radiol, 2011. 84(999): p. 251-64.
15. Wang, J. and X. Gu, *High-quality four-dimensional cone-beam CT by deforming prior images*. Phys Med Biol, 2013. 58(2): p. 231-46.
16. Zhang, H., M. Kruis, and J.J. Sonke, *Directional sinogram interpolation for motion weighted 4D cone-beam CT reconstruction*. Phys Med Biol, 2017. 62(6): p. 2254-2275.
17. Zhang, H., et al., *High quality 4D cone-beam CT reconstruction using motion-compensated total variation regularization*. Phys Med Biol, 2017. 62(8): p. 3313-3329.
18. Jarry, G., et al., *Characterization of scattered radiation in kV CBCT images using Monte Carlo simulations*. Med Phys, 2006. 33(11): p. 4320-9.
19. Ruhrnschopf, Ep, and K. Klingensbeck, *A general framework and review of scatter correction methods in cone beam CT. Part 2: scatter estimation approaches*. Med Phys, 2011. 38(9): p. 5186-99.
20. Stankovic, U., et al., *Improved image quality of cone beam CT scans for radiotherapy image guidance using fiber-interspaced antiscatter grid*. Med Phys, 2014. 41(6): p. 061910.
21. Bootsma, G.J., F. Verhaegen, and D.A. Jaffray, *Efficient scatter distribution esti-*

- tion and correction in CBCT using concurrent Monte Carlo fitting. *Med Phys*, 2015. **42**(1): p. 54-68.
22. Stankovic, U., et al., *Optimal combination of anti-scatter grids and software correction for CBCT imaging*. *Med Phys*, 2017. **44**(9): p. 4437-4451.
 23. Maier, J., et al., *Deep Scatter Estimation (DSE): Accurate Real-Time Scatter Estimation for X-Ray CT Using a Deep Convolutional Neural Network*. *Journal of Nondestructive Evaluation*, 2018. **37**(3).
 24. Kida, S., et al., *Cone Beam Computed Tomography Image Quality Improvement Using a Deep Convolutional Neural Network*. *Cureus*, 2018. **10**(4): p. e2548.
 25. Hansen, D.C., et al., *ScatterNet: A convolutional neural network for cone-beam CT intensity correction*. *Med Phys*, 2018. **45**(11): p. 4916-4926.
 26. Landry, G., et al., *Comparing Unet training with three different datasets to correct CBCT images for prostate radiotherapy dose calculations*. *Phys Med Biol*, 2019. **64**(3): p. 035011.
 27. Nomura, Y., et al., *Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network*. *Med Phys*, 2019. **46**(7): p. 3142-3155.
 28. Lee, H. and J. Lee, *A Deep Learning-Based Scatter Correction of Simulated X-ray Images*. *Electronics*, 2019. **8**(9).
 29. Chollet, F. *Keras*. 2015 [cited 2019; Available from: <https://github.com/keras-team/keras>].
 30. Kingma, D. and J. Lei, *Adam: A method for stochastic optimization*. *arXiv*, 2015.
 31. Zhang, H., et al., *mixup: beyond empirical risk minimization*. 2018.
 32. Agostinelli, S., et al., *Geant4—a simulation toolkit*. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2003. **506**(3): p. 250-303.
 33. Poludniowski, G., et al., *SpekCalc: a program to calculate photon spectra from tungsten anode x-ray tubes*. *Phys Med Biol*, 2009. **54**(19): p. N433-8.
 34. Poludniowski, G.G., *Calculation of x-ray spectra emerging from an x-ray tube. Part II. X-ray production and filtration in x-ray targets*. *Med Phys*, 2007. **34**(6): p. 2175-86.
 35. Poludniowski, G.G. and P.M. Evans, *Calculation of x-ray spectra emerging from an x-ray tube. Part I. electron penetration characteristics in x-ray targets*. *Med Phys*, 2007. **34**(6): p. 2164-74.
 36. Raudaschl, P.F., et al., *Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015*. *Med Phys*, 2017. **44**(5): p. 2020-2036.
 37. Suri, R.E., et al., *Comparison of scatter correction methods for CBCT*, in *Medical Imaging 2006: Physics of Medical Imaging*. 2006.
 38. Wang, Z., et al., *Image Quality Assessment: From Error Visibility to Structural Similarity*. *IEEE Transactions on Image Processing*, 2004. **13**(4): p. 600-612.
 39. Klein, S., et al., *elastix: a toolbox for intensity-based medical image registration*. *IEEE Trans Med Imaging*, 2010. **29**(1): p. 196-205.
 40. Ronneberger, O., P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015. p. 234-241.
 41. Arai, K., et al., *Feasibility of CBCT-based proton dose calculation using a histogram-matching algorithm in proton beam therapy*. *Phys Med*, 2017. **33**: p. 68-76.
 42. Park, Y.K., et al., *Proton dose calculation on scatter-corrected CBCT image: Feasibility study for adaptive proton therapy*. *Med Phys*, 2015. **42**(8): p. 4449-59.
 43. Vilches-Freixas, G., et al., *Comparison of projection- and image-based methods for proton stopping power estimation*

- using dual energy CT*. Physics and Imaging in Radiation Oncology, 2017. **3**: p. 28-36.
44. Ballabriga, R., et al., *Medipix3: A 64k pixel detector readout chip working in single photon counting mode with improved spectrometric performance*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2011. **633**: p. S15-S18.
45. Procz, S., et al., *X-ray and gamma imaging with Medipix and Timepix detectors in medical research*. Radiation Measurements, 2019. **127**.

Supplementary Materials

A. Scatter size prediction size determination

The network architecture was built to predict scatter distributions on a 32x32 image grid. As illustration for one patient projection (Figure S1), the bi-cubic downsampling and the bi-cubic upsampling did not show information loss after rebinning the scatter distribution to a 32x32 image grid. The mean absolute error between the original and the resized signal was 0.035%. Smaller image grids (i.e. 16x16) were investigated but led to larger mean absolute errors. This methodology was validated on multiple patient projections.

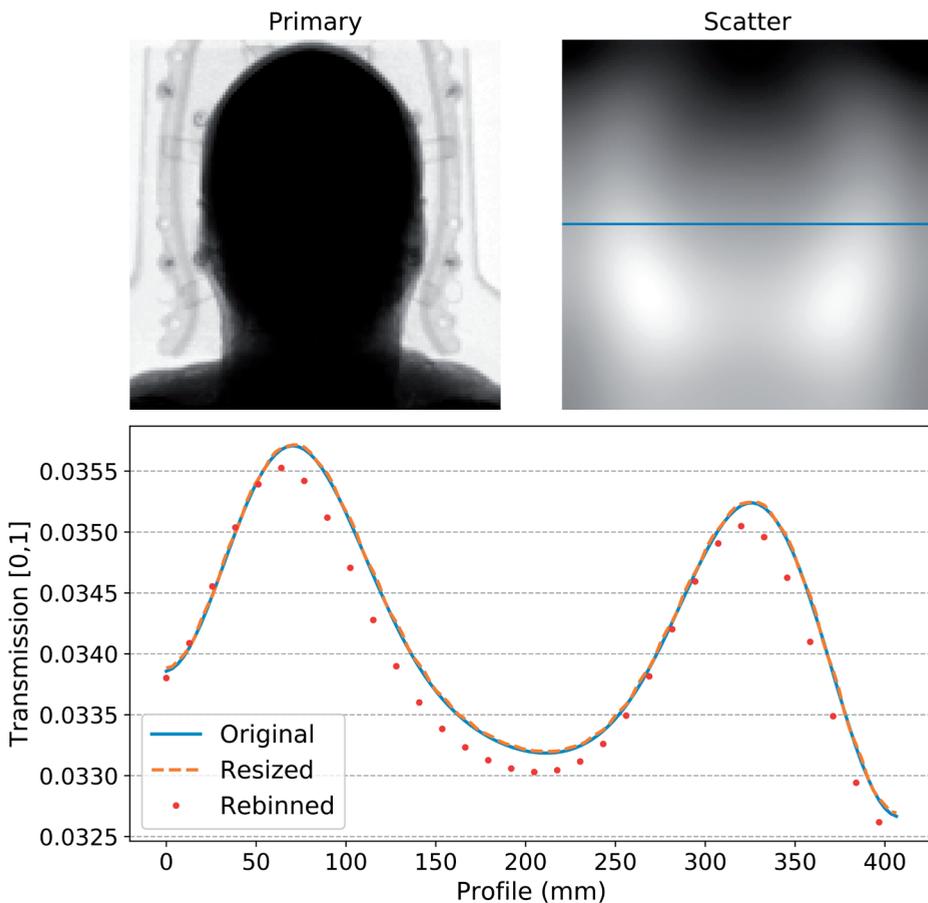


Figure S1. The top panels depict a simulated primary and a scatter fluence distribution for a single projection image. On projection level, a horizontal cross profile through the scatter fluence was plotted in the bottom panel for the original simulated scatter distribution (blue line), the rebinned scatter distribution on a 32x32 projection grid (red dots), and the upsampled scatter distribution to the original image grid (orange dashes).

B. Image Post-Processing

In the scatter correction workflow, the scatter is first predicted and then subtracted from the scatter sensitive input projection. The presence of non-natural materials in the patient, such as dental fillings, sometimes causes negative transmission values which are physically incorrect. **Figure S2** shows regions of negative pixel regions on a DCAE corrected X-ray projection with increased opacity. Post-processing after scatter subtraction was applied on projection level p according to **Equation S1** to correct for the presence of these negative pixel values after scatter subtraction.

$$p(p < T) = T \cdot e^{\frac{p-T}{T}}, \text{ where } T = 0.015 \quad (\text{Eq. S1})$$

The value of constant T in **Equation 1** was determined empirically in a range of transmission values between 0.00 and 0.05. The lowest overall error rate compared to the primary fluence simulation was observed with T equal to 0.015.

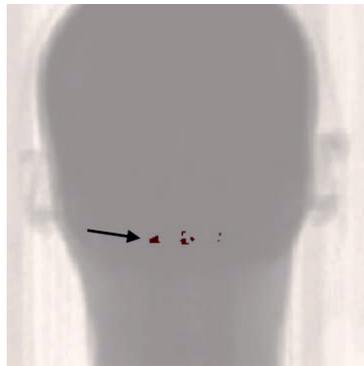


Figure S2. Scatter corrected X-ray projection from the MC test dataset. No post-processing was applied yet in this example, and the pixel regions with negative transmission values are indicated by the black arrow.

C. Flat panel detector response

Simulating the particle transport of the secondary electrons and optical photons in the active layer of the XRD 1642 Perkin Elmer flat-panel detector requires too much computing power in full CBCT trajectory simulations. Therefore, a separate Monte Carlo simulation module was developed to determine the detector energy response curve of the detector. Technical drawings of the flat-panel detector were used to calculate the detector energy response of the detector. In this separate module, monoenergetic photon pencil beams (10E6 histories per beam) were impinging perpendicularly on the detector in energy steps of 1 keV ranging from 1 keV up to 160 keV. For every pencil beam, the incoming particle energy and the average particle energy deposit were scored in the active layer of the detector. Then, the ratio of the average

deposited energy and the incoming energy was used to calculate the detector energy response curve (Figure S3).

In the main ImagingRing™ system on rails Monte Carlo model, the detector was simulated with an assigned Geant4 sensitive detector, which forces all incident photons to stop in the detector elements and requests their deposited kinetic energy. Based on the deposited kinetic energy of the incident photons, the sensitive detector will sample uniformly from the detector energy response curve (Figure S3) to determine whether a single photon is detected or ignored.

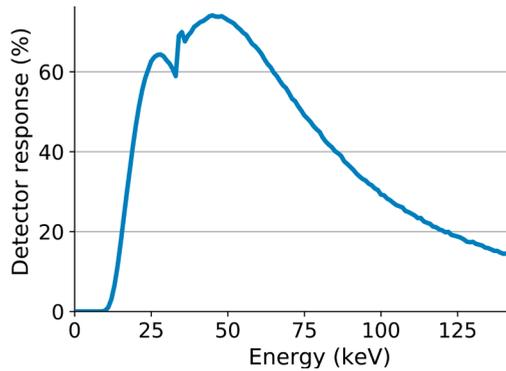


Figure S3. The simulated detector energy response curve of the XRD 1642 Perkin Elmer flat panel detector.

D. Heuristic scatter removal approach

The heuristic scatter correction approach removes object scatter on the projection level using a method based on the Scatter Kernel Method, described by Suri *et al.* (2006). The specific implementation models the scatter as two scatter components. The first scatter contribution S_S arises from modeling a single scatter interaction and still contains a reasonable amount of local information. The second contribution S_M arises from multiple scattering events and contains little local information. The basic principle of this correction is the computation of both scatter estimations based on the original input transmission images (T'). The sum of the two scatter components is then subtracted from T' to determine the corrected transmission image (T). The scatter estimations S_S and S_M are computed by deriving the point scatter estimations PS_S and PS_M , by mapping the original transmission image T' through a one-dimensional lookup-table (LUT). The scatter estimations S_S and S_M are obtained by blurring the point scatter estimations with a Gaussian kernel. The LUT describes a mapping from input transmission values to transmission values that represents either the forward or the isotropic scatter portion of the input transmission:

$$T = T' - (S_S(T') + S_M(T')) \quad PS_X(T') = LUT_X(T')$$

The lookup-tables (LUT_S and LUT_M) were determined experimentally. Solid water slabs with varying thicknesses 0cm and 40cm were placed with a specific distance (SD) to the detector and were sequentially imaged by a pencil beam. A projection image with a given solid water thickness delivered two measuring points, one for every scatter component in function of transmission values. Besides modelling the absolute scatter values via the LUTs, the scatter blurring was parametrized by fitting two Gaussian kernels into the measured line profiles, configuring them with σ_S and σ_M . Subsequently, each of the components, $S_S(T')$ and $S_M(T')$, is calculated by:

$$S_X(T') = \left(PS_X(T') * \left(\frac{1}{\sqrt{2\pi} \sigma_X} e^{-\frac{(3\sigma_X)^2}{2\sigma_X^2}} \right) \right) \frac{SD}{ODD}, \quad \text{where } x = [S, M]$$

As the scatter estimates are based on measurements made on a specific distance (SD), for the non-isotropic imaging setup, a subsequent mapping of the computed scatter onto the detector plane is performed by considering the object detector distance (ODD). The SD is defined by the ODD of the solid water phantom in the acquisition setup to measure LUT-data.

8

Chapter 8

Automatic multi-atlas based organs-at-risk segmentation in mice

Brent van der Heyden, Mark Podesta, Daniëlle BP Eekers, Ana Vaniqui,
Isabel P Almeida, Lotte EJR Schyns, Stefan J van Hoof, and Frank Verhaegen

Published in

British Journal of Radiology (2018)

<https://doi.org/10.1259/bjr.20180364>

Abstract

Background and Purpose: During the treatment planning of a preclinical small animal irradiation, which has time limitations for reasons of animal wellbeing and workflow efficiency, the time-consuming organ at risk (OAR) delineation is performed manually. This work aimed to develop, demonstrate, and quantitatively evaluate an automated contouring method for six OARs in a preclinical irradiation treatment workflow.

Materials and Methods: Micro cone-beam CT images of nine healthy mice were contoured with an in-house developed multi-atlas (MA) based image segmentation algorithm for six OARs: kidneys, eyes, heart, and brain. The automatic contouring was compared with the manual delineation using three quantitative metrics: the Dice Similarity Coefficient (DSC), 95th percentile Hausdorff Distance (HD), and the center of mass displacement (Δ COM).

Results: A good agreement between manual and automatic contouring was found for OARs with sharp organ boundaries. For the brain and the heart, the median DSC was larger than 0.94, the median 95th HD smaller than 0.44 mm, and the median Δ COM smaller than 1.93 mm. The absence of contrast medium resulted in lower values for the other OARs, but the median DSC was still larger than 0.74 for the left eye, 0.69 for the right eye, 0.89 for the left kidney and 0.80 for the right kidney.

Conclusion: The MA based segmentation algorithm was able to delineate six OARs with a relatively high accuracy. Segmenting OARs with sharp organ boundaries performed better than low contrast OARs. Our MA based algorithm is evaluated and demonstrated in a preclinical small animal irradiation research workflow.

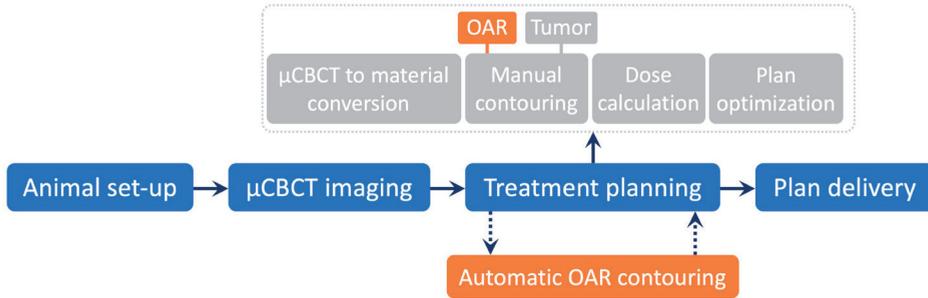


Figure 1. Preclinical small animal image-guided precision irradiation workflow. In the treatment planning system, the time consuming manual contouring of the OARs, often performed by biologists and not delineation experts, could be replaced by automatic OAR contouring (orange), while simultaneously other (grey) steps such as the CBCT to material conversion, the manual tumor delineation, the dose calculation, and the plan optimization are performed. (Abbreviations: OAR, organs-at-risk).

Introduction

A modern preclinical small animal irradiation workflow, using image-guided precision irradiation cabinets, consists of several steps [1] (Figure 1) that must fit into a tight time schedule since the animal is physically constrained under deep anesthesia. In the first steps of this preclinical workflow, the animal is anesthetized and imaged using the onboard micro cone-beam computed tomography (μ CBCT) scanner, which is frequently implemented in these irradiators. After finishing the μ CBCT image reconstruction, the images are loaded in a dedicated preclinical treatment planning system to design a quantitative treatment plan to deliver, following the pre-clinical irradiation workflow. In these treatment planning systems [2], such as SmART-Plan (Precision X-ray Inc., North Branford, CT) [3], an interactive tissue segmentation is performed based on a CT number histogram, and is followed by a time consuming manual delineation of both the target (e.g. the tumor) and the organs at risk (OARs). In the final step of the planning workflow, the irradiation treatment is optimized; the dose calculation of the irradiation plan is performed and evaluated quantitatively considering the dose volume histogram parameters. Potentially also other forms of imaging, such as bioluminescence photon imaging, may be performed while the animal is under anesthesia [4]. All these workflow steps should not require more than 20-60 minutes, for reasons of animal wellbeing and workflow efficiency [5].

The manual delineation of the OARs before plan delivery is one of the most time-consuming tasks, especially due to the high number of axial slices resulting from the sub-millimetric voxel sizes used in pre-clinical μ CBCT imaging. Much

effort was made to automatically segment OARs for human radiation treatment planning. E.g., open source packages for automatic image segmentation were developed [6], automatic contouring competitions were organized [7], and researchers considered the future perspectives of automatic contouring in radiotherapy [8].

In the literature, only a few efforts have been made to segment organ regions on images of rodents. Scheenstra *et al.* [9] presented an automated hybrid method to segment in vivo and ex vivo mouse brain magnetic resonance (MR) images that combined affine image registration and clustering techniques. Uberti *et al.* [10] applied a semi-automatic algorithm to extract the brain from in vivo mouse MR images by user-defined constraints. To improve the preclinical functional imaging workflow, Lancelot *et al.* [11] evaluated a multi-atlas based rat MR image segmentation algorithm and investigated whether the algorithm was capable of automatically extracting [¹⁸F]FDG radioactivity from volumes of interest of positron emission tomography (PET) images in the brain.

Other researchers investigated the feasibility of automatic image segmentation using contrast-enhanced fan-beam μ CT images. Yan *et al.* [12] developed a super-voxel based machine learning algorithm that used knowledge of the different time points in a dynamic contrast enhanced fan-beam μ CT scan of mice to classify super-voxels in multiple organ categories. Baiker *et al.* [13] presented a fully automated atlas based whole-body segmentation from low-contrast fan-beam μ CT images using a 4D digital mouse phantom [14] (MOBY) and a hierarchical atlas tree description. Most studies in (semi-) automatic tissue segmentation of small animal organs (regions) were for MR imaging and contrast-enhanced fan-beam μ CT. To our knowledge there are no studies investigating the feasibility of automatic OAR segmentation in a preclinical irradiation workflow using μ CBCT images. In this chapter, we describe an in-house developed multi-atlas (MA) based image segmentation algorithm and investigate whether automatic OAR segmentation is feasible using a pre-clinical in vivo non-contrast-enhanced μ CBCT dataset of healthy mice. In MA based algorithms, the atlas is referring to a μ CBCT dataset with already delineated OARs by a medical expert.

Materials and Methods

Mice dataset

Nine healthy female NMRI nude mice without tumors served as the control group in a biological study were μ CBCT imaged with an X-RAD 225Cx small animal radiation research system (Precision X-ray Inc., North Branford, CT) at an X-ray tube potential of 50 kVp, X-ray tube current of 5.59 mA, and a gantry rotation time of 0.5 rotations per minute [15]. During the μ CBCT imaging, all mice were under deep

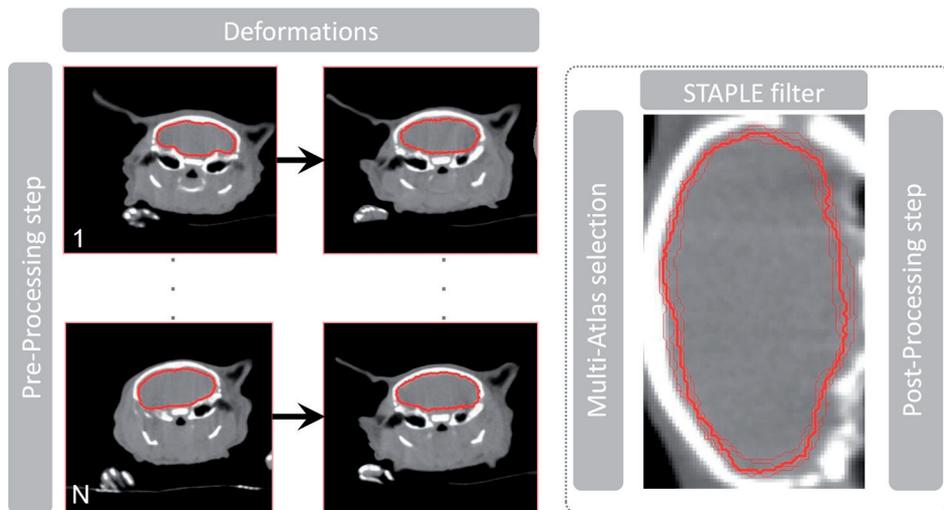


Figure 2. Summary of our in-house developed MA based image segmentation algorithm. A pre-processing step is completed before every atlas in the atlas database (1,2, ..., N atlases) is deformed to the unsegmented mouse (in this example in the neurological region). At the organ level, the atlases with the highest normalized cross-correlation (NCC) coefficients are selected. The Simultaneous Truth And Performance Level Estimation (STAPLE) filter algorithm calculates a consensus contour (thick contour) from the individual deformed mouse contours (thin lines) and the post-processing algorithm is executed in a last step (the right image is rotated 90 degrees for viewing purposes).

anesthesia i.e. isoflurane. The corrected projection data (flood field, dark field, and dead pixels) is reconstructed using the Feldkamp-Davis-Kress (FDK) back projection algorithm implemented in the open-source Reconstruction Toolkit [16] (RTK) in a $512 \times 512 \times 1024$ image matrix with a voxel dimension of $100 \times 100 \times 100 \mu\text{m}^3$. All nine mice were imaged at a 30 cGy imaging dose without contrast agent and in feet first-prone position including the nose cone, which is used to administer the anaesthesia gas. The imaging dose level was verified using a PTW TN300012 Farmer-type ionization chamber (PTW, Freiburg, Germany) according to the AAPM TG-61 protocol [17]. For each individual mouse, a medical doctor, experienced in mouse OAR delineations, contoured six OARs: the left kidney, right kidney, heart, brain, left eye, and the right eye. In this study, we assume these manually delineated contours as the ground truth.

In the pre-processing step, the reconstructed $512 \times 512 \times 1024$ image matrix is cropped to a $300 \times 300 \times 700$ image matrix. After cropping, the μCBCT image data is denoised using the multi-threaded curvature flow image filter (number of iterations = 5; time step = 0.05) implemented in the Insight Segmentation and Registration Toolkit (ITK). This edge-preserving denoising algorithm is smoothing the image to

obtain a better deformation in the last stage of the multistage image deformation algorithm explained in the next section.

In the deformation step, Elastix [18] and its Transformix submodule are used to apply a multistage deformable image registration (DIR). This DIR algorithm first calculates an affine transformation followed by a B-spline transformation. Here, the affine transformation eliminates the animal positioning dependency in the automatic OAR segmentation procedure. Considering the leave-one-out approach, the calculated deformation field is then used to deform the original μ CBCT and segmentation dataset from all mice in the remaining mouse atlas database ($N=8$) to the unsegmented mouse that is left out. A detailed list of the selected deformation parameters of the affine transformation and B-spline transformation are listed in the [Supplementary Material A](#).

After deformation of the mouse atlas dataset to the mouse that has been omitted, eight deformed μ CBCT and eight deformed segmentation datasets could be used as input for the STAPLE filter [19]. However, the MA based algorithm performs an atlas selection for every organ. I.e. different mouse atlases could be used to contour different OARs, all depending on the performance of the deformation in the region around the OAR. To select the best mouse atlas for every OAR, the eight deformed three-dimensional (3D) manually segmented OAR datasets are combined into a logical four-dimensional (4D) image vector, where the fourth dimension represents the number of deformed atlases (9 minus 1 in this study). The fourth dimension is summed logically to calculate a 3D bounding box that perfectly encloses the deformed OAR in the eight deformed mouse μ CBCT datasets. This 3D bounding box plus a small expansion, calculated for every OAR, is used to crop the deformed mouse μ CBCT dataset and the μ CBCT image data of the omitted unsegmented mouse.

In the last step of the atlas selection, for every organ, the normalized cross-correlation (NCC) coefficient is calculated between eight ($j=1, \dots, 8$) deformed cropped mouse atlases (A_j) and the cropped unsegmented mouse image data (M), where $A_{j,i}$ and M_i represent the voxel intensities at voxel index i . \bar{A}_j and \bar{M} are the mean voxel intensities in the μ CBCT images A and M . The NCC coefficient is equal to one when both μ CBCT datasets correlate perfectly and is equal to zero when no correlation is found.

$$NCC_j(A_j, M) = \frac{\sum_i (A_{j,i} - \bar{A}_j)(M_i - \bar{M})}{\sqrt{\sum_i (A_{j,i} - \bar{A}_j)^2} \sqrt{\sum_i (M_i - \bar{M})^2}}$$

NCC_j coefficients, calculated for each deformed atlas j , are sorted in descending order for every OAR segmentation. The structure deformations, corresponding to image deformations with the highest NCC_j coefficients, are used as input for the STAPLE algorithm, which calculates an OAR probability distribution. In our MA

based algorithm, the STAPLE filter implementation from ITK is adopted. After thresholding the OAR probability distribution, a raw automatic segmentation can be derived which is then morphologically smoothed in the post-processing step to obtain the final contour of the OAR.

The MA based image segmentation algorithm's pre-processing step, applying the STAPLE filter for all OARs, and the post-processing step were run on a computer with two Intel® Xeon® E5-2650 v3 (2.30 GHz) processors. All Elastix and Transformix jobs were submitted to an HTCCondor CPU cluster (144 cores, at least 2GB RAM per core) connected with a one gigabit data transfer Ethernet connection.

Segmentation evaluation metrics

For evaluation of all nine automatic OAR segmentations, the Dice Similarity Coefficient (DSC), the 95th percentile Hausdorff Distance (95th HD), and the center of mass displacement (Δ COM) are employed.

These three quantitative evaluation metrics are calculated between the ground truth OAR segmentation by the medical doctor (S_{MD}) and the automatic OAR segmentations from the MA based algorithm (S_{auto}).

1. The DSC calculates the overlap of two contours and is equal to one if both contours perfectly overlap.

$$DSC(S_{MD}, S_{auto}) = \frac{2|S_{MD} \cap S_{auto}|}{|S_{MD}| + |S_{auto}|}$$

2. The 95th HD is a measure to describe the 95th percentile boundary distance between two contours, where the best outcome is 0 mm, and is quantified as follows:

$$HD(S_{MD}, S_{auto}) = 95^{th} \text{ percentile} \left(\max_{md \in S_{MD}} \left\{ \min_{auto \in S_{auto}} \|md - auto\| \right\} \right)$$

3. The center of mass displacement (Δ COM) between the center of mass of two contours is calculated in a Cartesian coordinate system (x, y, z). With the Δ COM metric equal to zero, the mass centers of both contours perfectly overlap.

Results

Figure 3 reports the DSC, 95th HD, and Δ COM quantitative metrics for all six OARs. The automatically generated contour by the MA based image segmentation algorithm and the manual delineation, plotted on the μ CBCT image, are compared in **Figure 4** for each individual OAR and for all nine mice. The axial μ CBCT slices

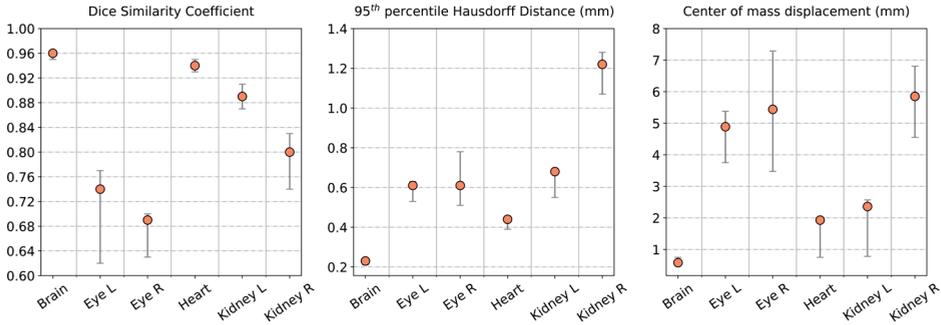


Figure 3. Quantitative evaluation of the automatic OAR segmentations in nine mice adopting the leave-one-out technique. The median evaluation metric value together with the 25th and the 75th percentiles whiskers are plotted.

shown in [Figure 4](#) were randomly chosen around the middle of the manually delineated OAR.

The average MA based auto-segmentation time of one mouse was 12 minutes and 40 seconds. From this segmentation time, 45% of the calculation time was required to write, transfer, read, and deform the mouse atlas dataset (#8), and 55% of the calculation time was needed to apply the post-processing step for six OARs. Only an average contouring time is reported in this study because the deformation calculations were submitted randomly to the HTCCondor CPU cluster, which is composed of computers with different processor units.

Discussion

The in-house developed MA based image segmentation algorithm performed better for structures with sharp boundaries such as the skull (i.e. brain) or the lungs (i.e. heart). For both structures the median DSC was larger than 0.94, the median 95th HD smaller than 0.44 mm, and the median ΔCoM smaller than 1.93 mm. In [Figure 3](#), a small spread, indicated by the 25th and 75th percentile whiskers, on the nine individual DSC, 95th HD, and ΔCoM metrics of the brain and the heart, is noticeable.

OARs having a lower image contrast are more difficult to contour in a non-contrast enhanced μ CBCT image. This difficulty applies to both the manual contouring (larger intra-observer variability) and the automatic contouring, which relies on manual contoured atlases. The absence of image contrast results in lower median DSCs than the brain or the heart, but still larger than 0.74 for the left eye, 0.69 for the right eye, 0.89 for the left kidney and 0.80 for the right kidney. The median 95th HD and ΔCoM are smaller for organs with sharp boundaries and larger for organs

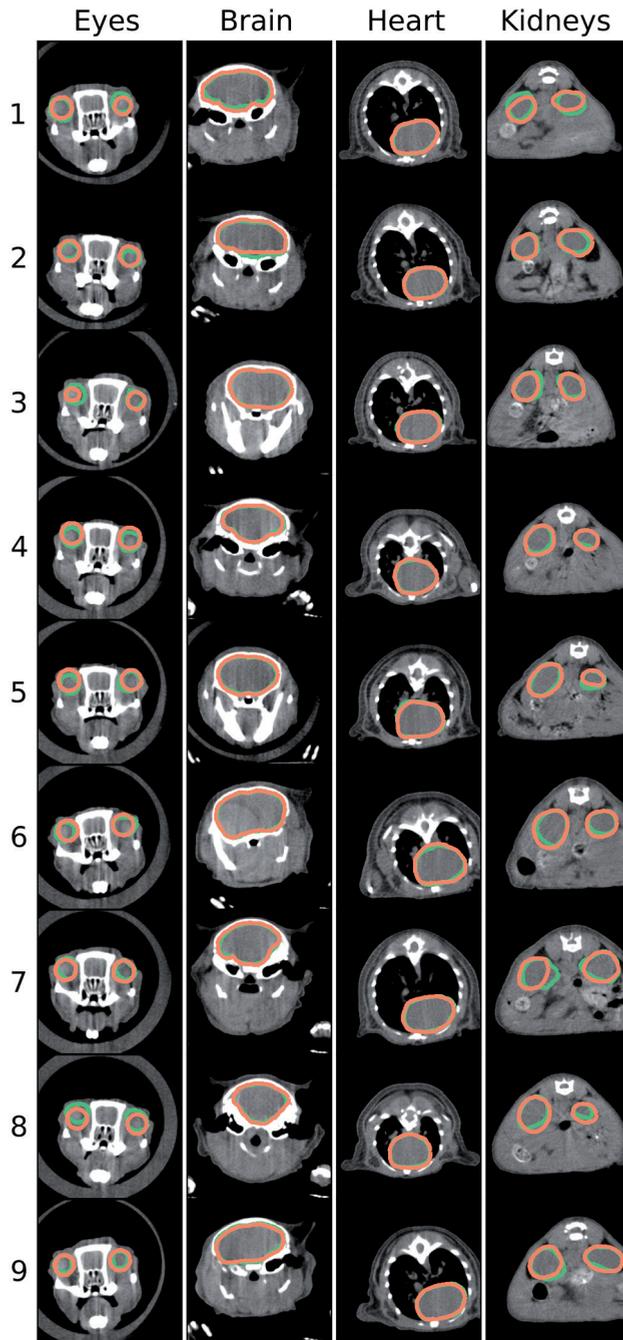


Figure 4. The automatic generated contour (green) and the ground truth delineation (orange) depicted for all nine mice (rows). The axial slices were randomly around the middle of the manually delineated OAR.

located in a region with lower tissue contrast. In particular, a larger median ΔCOM was obtained for the eyes and the kidneys, more specifically: 4.9 mm for the left eye, 5.4 mm for the right eye, 2.4 mm for the left kidney, and 5.9 mm for the right kidney. We expect the ΔCOM differences between the left kidney and the right kidney are due to the intra-observer variability resulting from the less distinguishable OAR boundary.

Studies such as Baiker *et al.* [13] and Wang *et al.* [20] reported DSCs of organs that were also included in our study. Baiker *et al.* [13], using fan-beam μCT image data, reported mean DSCs (± 1 standard deviation) to evaluate the segmentation accuracy of the kidneys (0.47 ± 0.08), the heart (0.50 ± 0.06)⁵, and the brain (0.73 ± 0.04). After applying the same statistical analyses in our study, using medians and percentiles, as in Baiker *et al.* [13], using means and standard deviations, our method obtained mean DSCs of 0.82 ± 0.09 , 0.94 ± 0.02 , and 0.96 ± 0.01 for the kidneys, heart, and brain, respectively.

The best DSC results reported in Wang *et al.* [20], using contrast-enhanced fan-beam μCT image data, were 0.91 for the heart, 0.71 for the left kidney, and 0.81 for the right kidney. These DSCs are closer, but still lower than the results presented in our study. However, our MA based algorithm only used eight mouse atlases to contour a mouse with lacking OAR segmentations. In addition, acquiring images in the X-RAD 225Cx small animal radiation research system can vary from 40 seconds up to 10 minutes, depending on desired resolution and number of projections.

Our non-contrast-enhanced μCBCT method, imaging the mice for about 2 minutes, is susceptible to respiratory and cardiac motion. Animal movement is less relevant for imaging neurological regions but is of higher relevance while performing cardio-thoracic or abdominal imaging. In the μCBCT image reconstruction, animal motion causes blurred organ edges, which are not beneficial for manual contouring and our MA based algorithm. In this study, no contrast agent was injected in the mice before imaging. Because contrast agent materials consist of elements with a higher atomic number, this will result in more X-ray attenuation and therefore in sharper edges around organs that incorporate contrast agent. Taking into account the fact that the multistage DIR algorithm uses the mutual information metric, which assumes a relation between the probability distributions of the intensities of the atlas μCBCT and the unsegmented mouse μCBCT , our MA based algorithm is also applicable to contrast-enhanced μCBCT image acquisitions. Applying the MA based algorithm on contrast-enhanced image data could result in much better OAR contours resulting from the improved image contrast between organ boundaries. In small animal body regions that are susceptible to respiratory motion, the combination of contrast-enhanced μCBCT and 4D μCBCT imaging can be beneficial for the

5 Value calculated with DataThief 3 (<https://datathief.org/>)

image contrast and thus the automatic image segmentation, mainly because an anesthetized animal (i.e. isoflurane) breathes with short in- and exhale time periods and much more longer rest phases, e.g. $\pm 0.6s$ and $\pm 2.2s$ respectively [21]. We expect that the μ CBCT motion artefacts caused by cardiac motion will not be improved by contrast-enhanced 4D μ CBCT because the periodically fast heart-beating rate of a mouse is only causing small deformations in the mouse anatomy.

To implement the MA based image segmentation method in the preclinical small animal image-guided precision irradiation workflow, the method must be accurate, and the calculation time must be kept as low as possible. An accurate segmentation prevents that the operator of the treatment planning system has to correct the contour afterwards, which will increase the planning time. An average calculation time of 12 minutes and 40 seconds was required to perform all pre-processing, deformation, STAPLE filter, and post-processing for one mouse. These three steps include the edge preserving smoothing, the deformations and transformations of a $300 \times 300 \times 700$ image matrix, and include the post-processing of six OARs. In these ~ 13 minutes, the machine operator can simultaneously perform the tissue assignment needed by the dose engine, tumor delineation, target location, and beam planning, as illustrated in Figure 1.

In human OAR image segmentation, many anatomical degrees of freedom are involved, such as; large anatomical changes due to surgery or a medical disease, gender, patient weight, patient size, patient age, and metal artefacts (e.g. dental fillings, implants, etc.). The MA based segmentation algorithm is expected to perform better on a mice image dataset because these variations are more limited, since they mice are from the same strain, usually have the same age, eat the same food etc. However, it still needs to be investigated whether the MA based segmentation algorithm is capable to automatically contour different mouse models, although we expect the algorithm to perform well as long the anatomical variations are limited.

In this investigation, the automatic contouring of multiple OARs from different treatment sites was studied. In pre-clinical image-guided high precision small animal irradiators, the OARs are only delineated nearby the treatment site of interest because of the small millimetric irradiation beams, e.g. the brain and eyes for neurological treatments, the kidneys for pancreatic treatments, the heart for lung treatments, etc. Independent from the treatment site, the μ CBCT image data can be further cropped or reconstructed in smaller image volumes which results in improved calculation times in every step of our MA based algorithm. Furthermore, the MA based segmentation algorithm is developed in a way that the software can also be used to automatically contour other rodents and even humans, although we expect the algorithm to perform better on rodent image data because the larger anatomical differences in humans.

Conclusion

This chapter is the first study that addresses automatic OAR segmentation in mouse μ CBCT images using a MA based image segmentation algorithm. This in-house developed MA based algorithm is evaluated in a preclinical small animal irradiation workflow. The automatic contouring accuracy of the MA based algorithm was quantitatively evaluated with the DSC, the 95th HD, and the Δ COM for six different OARs: left kidney, right kidney, heart, brain, left eye, and right eye. Using a powerful computer cluster, the algorithm automatically contoured the six OARs in a relatively small time-interval. The OARs with sharp organ boundaries are contoured very accurately, and for the OARs with less sharp organ boundaries, the MA based image segmentation algorithm produces contours that are usable but may require some manual corrections.

References

1. Verhaegen, F., P. Granton, and E. Tryggestad, *Small animal radiotherapy research platforms*. Phys Med Biol, 2011. **56**(12): p. R55-83.
2. Verhaegen, F., et al., *A review of treatment planning for precision image-guided photon beam pre-clinical animal radiation studies*. Z Med Phys, 2014. **24**(4): p. 323-34.
3. van Hoof, S.J., P.V. Granton, and F. Verhaegen, *Development and validation of a treatment planning system for small animal radiotherapy: SmART-Plan*. Radiother Oncol, 2013. **109**(3): p. 361-6.
4. Yahyanejad, S., et al., *Complementary use of bioluminescence imaging and contrast-enhanced micro-computed tomography in an orthotopic brain tumor model*. Mol Imaging, 2014. **13**.
5. Balvert, M., et al., *A framework for inverse planning of beam-on times for 3D small animal radiotherapy using interactive multi-objective optimisation*. Phys Med Biol, 2015. **60**(14): p. 5681-98.
6. Zaffino, P., et al., *Technical Note: plastimatch mabs, an open source tool for automatic image segmentation*. Med Phys, 2016. **43**(9): p. 5155.
7. Raudaschl, P.F., et al., *Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015*. Med Phys, 2017. **44**(5): p. 2020-2036.
8. Sharp, G., et al., *Vision 20/20: perspectives on automated image segmentation for radiotherapy*. Med Phys, 2014. **41**(5): p. 050902.
9. Scheenstra, A.E., et al., *Automated segmentation of in vivo and ex vivo mouse brain magnetic resonance images*. Mol Imaging, 2009. **8**(1): p. 35-44.
10. Uberti, M.G., M.D. Boska, and Y. Liu, *A semi-automatic image segmentation method for extraction of brain volume from in vivo mouse head magnetic resonance imaging using Constraint Level Sets*. J Neurosci Methods, 2009. **179**(2): p. 338-44.
11. Lancelot, S., et al., *A multi-atlas based method for automated anatomical rat brain MRI segmentation and extraction of PET activity*. PLoS One, 2014. **9**(10): p. e109113.
12. Yan, D., et al., *A Novel Mouse Segmentation Method Based on Dynamic Contrast Enhanced Micro-CT Images*. PLoS One, 2017. **12**(1): p. e0169424.

13. Baiker, M., et al., *Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data*. Med Image Anal, 2010. **14**(6): p. 723-37.
14. Segars, W.P., et al., *Development of a 4-D digital mouse phantom for molecular imaging research*. Mol Imaging Biol, 2004. **6**(3): p. 149-59.
15. Verhaegen, F., et al., *ESTRO ACROP: Technology for precision small animal radiotherapy research: Optimal use and challenges*. Radiother Oncol, 2018. **126**(3): p. 471-478.
16. Rit, S., et al., *The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)*. J Phys Conf Series, 2014.
17. Ma, C.M., et al., *AAPM protocol for 40-300 kV X-ray beam dosimetry in radiotherapy and radiobiology*. Med Phys, 2001. **28**(6): p. 868-93.
18. Klein, S., et al., *elastix: a toolbox for intensity-based medical image registration*. IEEE Trans Med Imaging, 2010. **29**(1): p. 196-205.
19. Rohlfing, T., D.B. Russakoff, and C.R. Maurer, Jr., *Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation*. IEEE Trans Med Imaging, 2004. **23**(8): p. 983-94.
20. Wang, H., D.B. Stout, and A.F. Chatzioannou, *Estimation of mouse organ locations through registration of a statistical mouse atlas with micro-CT images*. IEEE Trans Med Imaging, 2012. **31**(1): p. 88-102.
21. van der Heyden, B., et al., *The influence of respiratory motion on dose delivery in a mouse lung tumour irradiation using the 4D MOBY phantom*. Br J Radiol, 2017. **90**(1069): p. 20160419.

Supplementary Material

A. Deformation parameters (Elastix)

Supplementary 1. Elastix and Transformix parameters used as input for the affine transformation and the B-spline transformation, respectively. In this table, a cross-symbol (x) is used to specify parameters that were not defined in the respective parameter file.

Parameter name	Affine transform	B-spline transform
Registration	MultiResolutionRegistration	
FixedImagePyramid	FixedRecursiveImagePyramid	
MovingImagePyramid	MovingRecursiveImagePyramid	
Interpolator	BSplineInterpolator	
Metric	AdvancedMattesMutualInformation	
Optimizer	AdaptiveStochasticGradientDescent	
ResampleInterpolator	FinalBSplineInterpolator	
Resampler	DefaultResampler	
NumberOfResolutions	3	
ImagePyramidSchedule	4 4 4 2 2 2 1 1 1	
FinalGridSpacingInVoxels	x	8 8 8
GridSpacingSchedule	x	4 4 4 2 2 2 1 1 1
AutomaticScalesEstimation	True	x
AutomaticTransformInitialization	True	x
HowToCombineTransforms	Compose	
MaximumNumberOfIterations	1000	1500
AutomaticParameterEstimation	True	x
UseAdaptiveStepSizes	True	
ErodeMask	False	
UseDirectionCosines	True	
ImageSampler	Random	
NumberOfSpatialSamples	10,000	15,000
NewSamplesEveryIteration	True	
UseRandomSampleRegion	True	
BSplineInterpolationOrder	2	
FinalBSplineInterpolationOrder	2	

9

Chapter 9

Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network

Brent van der Heyden*, Wouter RPH van de Worp*, Ardy van Helvoort, Jan Theys, Annemie MWJ Schols, Ramon CJ Langen, and Frank Verhaegen

* Equal contribution

Published in

Journal of Applied Physiology (2019)

<https://doi.org/10.1152/jappphysiol.00465.2019>

Abstract

Background and Purpose: The loss of skeletal muscle mass is recognized as a complication of several chronic diseases and is associated with increased mortality and a decreased quality of life. Relevant and reliable animal models in which muscle wasting can be monitored non-invasively over time are instrumental to investigate and develop new therapies. In this work, we developed a fully automatic deep learning algorithm for segmentation of micro cone beam CT (μ CBCT) images of the lower limb muscle complex in mice, and subsequent muscle mass calculation.

Materials and Methods: A deep learning algorithm was trained on manually segmented data from 32 mice. Muscle wet mass measurements were obtained of 47 mice and served as a dataset for model validation and the reverse model validation. The automatic algorithm performance was approximately 150 times faster than manual segmentation. Reverse validation of the algorithm showed high quantitative metrics substantiating the robustness and accuracy of the model, i.e. a Dice Similarity Coefficient of 0.93, a Hausdorff distance of 0.4 mm, and a center of mass displacement of 0.1 mm.

Results: A high correlation ($R^2=0.92$) was obtained between the CT-derived muscle mass measurements and the muscle wet masses. Longitudinal follow-up revealed time-dependent changes in muscle mass that separated control from lung tumor bearing mice, which was confirmed as cachexia.

Conclusion: In conclusion, this deep learning model for automated assessment of the lower limb muscle complex provides highly accurate non-invasive longitudinal evaluation of skeletal muscle mass. Furthermore, it facilitates the workflow and increases the amount of data derived from mouse studies, while reducing the animal numbers.

Conflict of Interest

The NVIDIA Quadro P6000 GPU used for this research was donated by the NVIDIA Corporation.

Acknowledgements

The authors would like to thank Miriam van Dijk (Danone Nutricia Research, the Netherlands) for her valuable contribution to this work.

Supplementary Material

Online repository: <https://figshare.com/s/4f1610aa2c6991056cfe>

Introduction

Skeletal muscle is the largest tissue in the body and comprises approximately 40% of the total body mass in normal weight humans [1]. It allows voluntary movement, including mobility and breathing, but also contributes to numerous metabolic processes including substrate disposal and thermogenesis [2, 3]. Skeletal muscle is one of the most dynamic and plastic tissues of the body, responding to changes in functional or metabolic demand, most obviously in response to increased or reduced physical activity resulting in either hypertrophy or atrophy [4]. However, these characteristics also enable skeletal muscle to be used as energy and protein reserve that can be deployed during periods of starvation or during acute and chronic catabolism as seen during trauma or infections and cancer cachexia leading to accelerated muscle loss [5].

Loss of skeletal muscle mass, especially in the locomotor muscles, is a recognized complication of several chronic diseases including chronic obstructive pulmonary disease, chronic heart failure, and chronic kidney disease and in most malignant cancers [6]. Muscle wasting is well known to be associated with increased mortality and decreased quality of life of patients [7-9]. As muscle wasting is an unmet medical need, the development of effective therapeutic strategies involves pre-clinical research, which requires reliable assessment of muscle mass in animal models.

In the clinical setting, several imaging modalities such as Computed Tomography (CT), Cone Beam CT (CBCT), dual-energy X-ray absorptiometry (DEXA), or magnetic resonance imaging (MRI) have been applied for years as practical and precise methods to assess body composition, and skeletal muscle volumes [10-18]. In contrast, skeletal muscle mass evaluation in pre-clinical studies is mostly based on assessment of muscle wet masses or muscle fiber cross sectional area by post-mortem histological examination. Both methods require muscle dissection and consequently are terminal experiments. This precludes longitudinal follow-up of muscle mass changes, and assessment of time-dependent changes in muscle mass therefore requires inclusion of multiple animals in the experimental design. Recently, novel approaches to non-invasively determine muscle volume, including changes over time of individual experimental animals using MRI and micro CT were described [19-23].

Micro CT is a dedicated X-ray based imaging technology for small animals that allows for tissue density quantifications, which makes micro CT a good image technique for longitudinal muscle mass assessment. To investigate muscle mass changes quantitatively, manual segmentation of the muscle volume on the CT reconstruction of the experimental animal is required, which in itself is a delicate and very time-consuming task for the researcher.

Many efforts have been made to develop automated CT segmentation techniques for patients, but surprisingly, there have only been sparse reports on fully automatic soft-tissue CT segmentation techniques in experimental animals. Baiker *et al* investigated an automatic reference atlas based whole body segmentation method in μ CT images using a digital mouse phantom [24]. More recently, our group investigated organ-at-risk segmentation in mice prior to irradiation with a multi-atlas image deformation-based approach [25]. Clearly, availability of automated muscle segmentation applications for mice would tremendously facilitate CT-based assessment of skeletal muscle mass in preclinical research. Moreover, it will strongly contribute to a reduction in the number of animals used for longitudinal studies that evaluate muscle mass changes, which is in line with the 3R principle embedded in national and international legislation and regulations on the protection of animals used for scientific purposes. To our knowledge, no literature exists on the fully automatic segmentation of skeletal muscle in mice. In this work we developed and validated a new method to automatically segment skeletal muscle in mice using artificial intelligence techniques which could be widely applied in animal models used to study muscle atrophy or hypertrophy.

Materials and Methods

A summarizing flowchart of the study set-up is shown in [Figure 1](#). Further details are described in the subsequent chapter sections.

Mouse experimental and imaging procedures

Animal preparation

This work was conducted in accordance with institutional guidelines for the care and use of laboratory animals established by the Ethics Committee for Animal Experimentation of the Maastricht University, in full compliance to national legislation on animal research following the European Directive 2010/63/EU for the use of animals for scientific purposes, and is part of a set of experiments to establish an orthotopic model of lung cancer cachexia. Male and female mice at 9-10 weeks (129S2/SvPasCrl, Charles River Laboratories) were socially housed in a climate-controlled room (12:12 dark-light cycle with a constant room temperature of $21\pm 1^\circ\text{C}$). Mice were given *ad libitum* access to food (AIN-93M) and water. After 1 week of acclimatization, mice were randomly allocated to either sham control ($n=15$) or tumor bearing (TB) group ($n=68$). All animals underwent surgery at the age of 12 weeks, at a standardized time window during their inactive period of the day. While anesthetized using a mixture of air and isoflurane (4% induction, 2% maintenance) and appropriate analgesia,

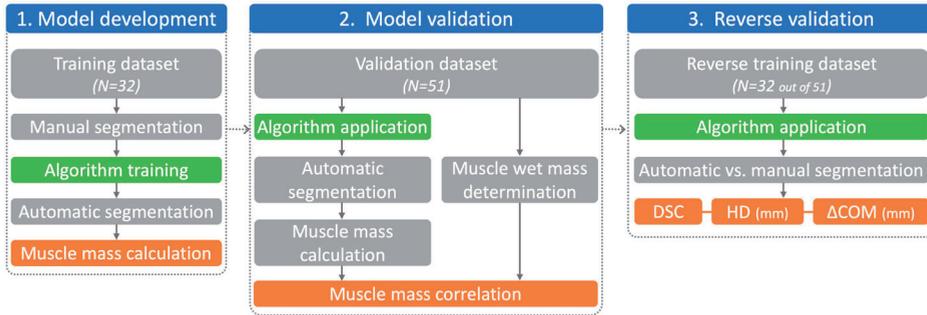


Figure 1. Study set-up flowchart. In the first step: 1. Model development, the model was trained on a manually segmented dataset of 32 mice. After training, the model segmented the lower limb muscle complex on micro cone beam CT (CBCT) scans and subsequently calculated the muscle mass using a pre-determined conversion curve. In step 2. Model validation was finally performed on a validation dataset of 47 mice (4 out of 51 mice were excluded). As validation, the muscle mass determined by the model was correlated to muscle wet mass measurements. In step 3. a reverse validation approach was adopted to calculate three quantitative metrics between the automatic and the manual muscle segmentations.

sham control mice received an intra-pulmonary injection with 15 μ l of matrix (Matrigel, Corning) and the TB mice (orthotopic lung tumor model) received an intra-pulmonary injection with 15 μ l of Matrigel containing lung epithelium-derived adenocarcinoma cells [26]. Mice were monitored daily. Based on a animal welfare scoring list, or upon a loss of body mass >20%, humane endpoints were applied. Furthermore, tumors were scored for characteristics, size and progression. When a total tumor volume of 500 mm³ was reached or in case of signs of dyspnea due to tumor growth, pulmonary constriction or metastasis, humane endpoints were applied.

Animal imaging

At baseline and weekly after surgery, μ CBCT imaging was performed for all mice to assess lung tumor development (not reported in this chapter) and detect muscle volume changes over time. For this, mice were anesthetized as outlined above, placed in prone position with toes facing the flanks (foot and tibia angle $\pm 90^\circ$), and scanned using a μ CBCT scanner (X-RAD 225Cx, Precision X-ray Inc., North Branford, USA) at an X-ray tube potential of 50 kVp, X-ray tube current of 5.6 mA, and an imaging time of 2 minutes [27]. The imaging dose of 30 cGy was verified using a PTW TN300012 Farmer-type ionization chamber (PTW, Freiburg, Germany) according to the AAPM TG-61 protocol [28]. The μ CBCT projection data was reconstructed using the pilot Feldkamp back projection algorithm with a voxel dimension of 100 x 100 x 100 μ m³ [29].

At the end of the experiment, mice were scanned and subsequently sacrificed using pentobarbital overdose to evaluate skeletal muscle mass. The soleus, plantaris, gastrocnemius, tibialis anterior (TA) and extensor digitorum longus (EDL) muscles including tendons were collected from both hind limbs, using standardized dissection methods with a precision coefficient of 0.93 ± 0.05 . Subsequently, muscles were immediately weighed in pairs on an analytical balance with precision of ± 0.1 mg and a linearity of 0.2 mg (CP64, Sartorius, Goettingen, Germany). 32 baseline μ CBCT scans of TB mice were used to train the newly developed algorithm and 47 mice (n=14 sham, n=33 TB) were used to validate the performance of the automatic muscle segmentation algorithm. The 47 mice were included from a group of 51 mice, of which 4 mice met the exclusion criteria (based on deviation of the expected mass-ratio of individual muscles, indicative of imperfections in the dissection procedure). Due to logistic reasons, the time interval between imaging and the muscle collection (day of sacrifice) of the validation group was either 2 days (n=24) or 0 days (n=23).

Automatic image segmentation

Deep learning algorithm

As training dataset for our algorithm, the lower limb muscle complex (including the musculus gastrocnemius, soleus, plantaris, TA and EDL) was manually segmented on the reconstructed cross sectional μ CBCT images that were acquired at the study baseline for a training dataset of 32 mice. The segmentation was performed by an experienced scientist with appropriate anatomical knowledge using the SmART-ATP software (Precision X-ray Inc., North Branford, CT) [30]. During the manual segmentation of the training dataset, the segmentation time was recorded to calculate the average manual segmentation time for one mouse. Images of five mice were randomly selected from the training dataset and were segmented twice with a time interval of 9 months between the segmentations to investigate the inter-observer variability or reproducibility of the manual segmentations.

The deep learning algorithm (two-step 3D U-Net model) was trained to segment the lower limb muscle complex volume on μ CBCT images (Hounsfield Units, HU) of mice [31, 32]. This 3D convolutional neural network architecture consisted of three encoding layers and three decoding layers, used weighted cross entropy as loss function and used a dropout ratio of 0.5. The neural network was applied in two successive steps and made use of Tensorflow (Python 2.7) in combination with the NVIDIA CUDA® Deep Neural Network library (cuDNN) computational kernels. The algorithm was executed on a NVIDIA Quadro P6000 (24 GB) graphics processing unit (GPU).

An automatic pre-processing step was performed on the μ CBCT mice dataset before the images were used by the deep learning algorithm. The μ CBCT volume was first cropped in the transversal plane to a [256, 256, Z] full resolution, where Z is the number of reconstructed μ CBCT slices. Next, the full resolution μ CBCT volume was resized with a cubic interpolation method to an X, Y, Z dimensions of [128, 128, Z/2]. A similar method, with nearest-neighbor interpolation, was applied on the manually segmented muscle volumes. In the final pre-processing step, both the down sampled and full resolution μ CBCT volumes were normalized between -400 HU and 1000 HU.

In the first step, a 3D U-Net model was trained (350 epochs) to segment the muscle volume on the down sampled and pre-processed μ CBCT images. The relative position of the automatically segmented muscle volume with respect to the down sampled μ CBCT image was then used to extract a volume-of-interest (VOI) with preset dimensions (64 x 64 x 192 pixels) in the full resolution μ CBCT dataset. In the second step (350 epochs), the extracted full resolution VOI was used to train the second 3D U-Net.

Reverse validation

To evaluate the robustness and the consistency of the automatic segmentation algorithm, a reverse validation approach was applied. First, the 'forward' automatic segmentation algorithm that was trained on the manually segmented dataset (32 mice) was applied on a validation dataset (51 mice). Here, the exclusion criterion was not adopted because the segmentations are not subject to experimental errors such as in the muscle wet mass experiments. In the reverse validation approach, a new model was trained based on the already automatically segmented muscle volumes. Here, the automatically segmented muscle volumes of 32 mice were randomly selected from the complete validation dataset (51 mice) to have a similar sized trained dataset compared to the original 'forward' segmentation algorithm. After training the 'reverse' model, this model was applied to segment the training dataset consisting of manual ground truth segmentations.

Finally, the automatic segmentation from the reverse algorithm was compared with the manual segmentation using three quantitative parameters, including the Dice similarity coefficient (DSC), 95th percentile Hausdorff distance (HD, in mm) and the center of mass displacement (Δ COM, in mm). The DSC indicates the volumetric overlap between the manual and the automatic muscle segmentation, the HD calculates the maximum distance of a point in the manual segmentation to the nearest point in the automatic segmentation considering the voxel dimensions, and the Δ COM calculates the 3D displacement of the mass centers of the manual and automatic segmentation in a Cartesian coordinate system.

μ CBCT to mass density conversion

A μ CBCT to mass density (CT2MD) conversion curve was required to calculate mass densities in every reconstructed μ CBCT voxel. This CT2MD curve was obtained by scanning a cylindrical mini-phantom (diameter = 3 cm, length = 1 cm) (SmART Scientific Solutions BV, Maastricht, the Netherlands) which is composed of a solid water bulk, two air inserts, and 10 tissue mimicking inserts of 3.5 mm diameter with known mass densities (e.g. adipose 0.95 g/cm³, soft tissue 1.06 g/cm³, and bone inserts 1.33-1.83 g/cm³). The mean CT numbers in HUs were calculated in circular regions-of-interest in the middle of the inserts and were then related to the certified mass densities of the cylindrical inserts in the form of a CT2MD conversion curve.

To determine muscle mass using μ CBCT imaging, the CT2MD conversion curve is applied on the reconstructed μ CBCT mice dataset (HUs) to obtain a three-dimensional density matrix (g/cm³). After applying the CT2MD curve on the reconstructed μ CBCT mice datasets, the binary masks resulting from our automatic muscle segmentation algorithm was applied on the converted mass density volume and the knowledge of the reconstructed voxel dimensions was used to calculate the muscle mass non-invasively. An image intensity-based thresholding technique was applied on the μ CBCT image dataset to remove bone from the analysis.

Results

Reverse validation and segmentation reproducibility

Three quantitative metrics (DSC, HD and Δ COM) \pm 1 standard deviation (SD) were calculated to evaluate the accuracy of the 'reverse' automatic segmentation algorithm. On average (\pm 1 SD), 158 \pm 8 axial μ CBCT slices were automatically segmented by the segmentation algorithm.

The DSC (\pm 1 SD) was equal to 0.93 \pm 0.03, the HD (\pm 1 SD) was equal to 0.4 \pm 0.2 mm and the Δ COM (\pm 1 SD) was equal to 0.1 \pm 0.1 mm indicating a good agreement compared to the manual ground truth muscle segmentations. To investigate the manual segmentation reproducibility of the human observer, these metrics were also calculated between two manual segmentations of five mice. Here, the DSC was equal to 0.95 \pm 0.01, the HD was equal to 0.4 \pm 0.1 mm, and the Δ COM was equal to 0.18 \pm 0.03 mm.

CT-derived muscle mass highly correlates with muscle wet masses

The combined muscle wet masses of the lower limb were determined in a validation set consisting of 47 mice and compared to the muscle masses calculated by the automatic segmentation algorithm (Figure 2). A trendline (blue) was fitted through all data points by linear regression ($R^2=0.92$). When separated for the time span between

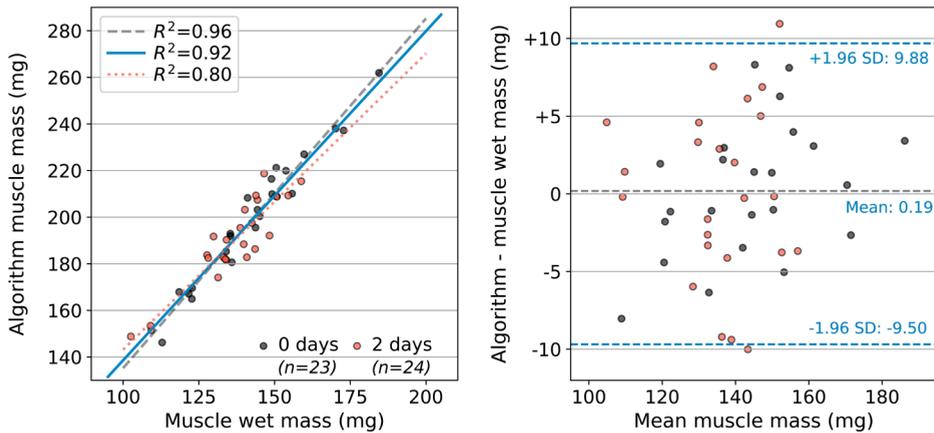


Figure 2. Left panel: linear relationships between the muscle wet mass measurements (in one leg) and the muscle masses determined by the automatic segmentation algorithm ($R^2=0.96$, $Y=1.50X - 14.85$; Subset weighted after 0 days: $R^2=0.92$, $Y=1.42X - 2.76$; Subset weighted after 2 days: $R^2=0.80$, $Y=1.27X + 16.04$). Right panel: Bland-Altman plot between the algorithm muscle mass and the experimental muscle wet mass.

μ CBCT scan and muscle wet mass assessment, the correlation further increased for the 23 mice for which muscle tissue was collected directly after imaging ($R^2=0.96$), compared to 24 mice of which muscle tissue was excised 2 days after imaging ($R^2=0.80$).

The ratio between the muscle wet mass and the algorithm-based muscle mass was a constant factor of 0.72 ± 0.03 and can be attributed to the excision of a select set of muscles from the hind limb musculature that can be excised with high accuracy. After factor correction of the algorithm-defined muscle mass, a Bland-Altman plot was constructed to analyze the agreement between the two methods (right panel of **Figure 2**). No proportional bias was found, meaning that the two methods agree equally through the range of measurements.

Longitudinal assessment of muscle mass

To study muscle mass changes over time and detect the onset of muscle wasting, longitudinal evaluation of muscle mass is essential. To evaluate the automatic follow-up capabilities of our algorithm for individual mice, muscle masses were assessed for repetitive analyses of 6 randomly selected mice (3 sham control mice and 3 TB mice) in the validation dataset. The muscle masses shown in **Figure 3** were calculated relative to the automatic measurement at baseline (day 0). These results show that the algorithm is capable of distinguishing muscle mass responses between sham control mice and TB mice developing muscle wasting.

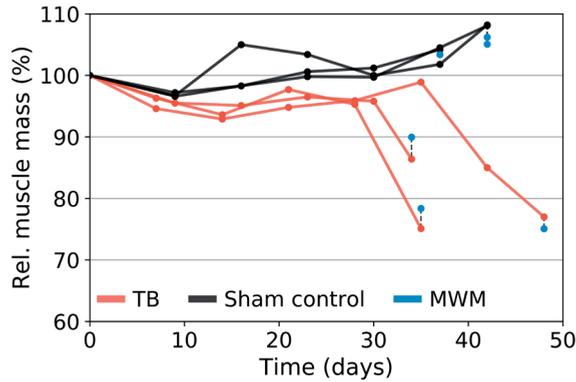


Figure 3. Follow-up curves determined by our automatic muscle segmentation algorithm for three tumor bearing (TB) mice (orange) and three sham control mice (black). The orange and black markers indicate the time of scanning, the lines that connect the markers are linearly interpolated. The blue markers present the experimental muscle wet mass (MWM) at the last timepoint after applying the fitted linear relationship between the MWM and the algorithm muscle mass. The absolute (calculated) muscle mass at the baseline measurement varied slightly (time=0 days: 220.3 ± 18.1 mg).

One TB mouse and one sham control mouse were randomly chosen to evaluate the automatic contouring with manual contouring. In Supplementary Material, animations in GIF-format are provided to visualize the automatic muscle segmentation and the manual muscle segmentation of these two mice in the axial, coronal and sagittal μ CBCT viewing planes.

Figure 4 shows the automatic muscle segmentation (orange contour) and the manual muscle segmentation (yellow contour) on the μ CBCT scan of these mice at the baseline measurement and at the last timepoint of the study. Because the mice were scanned at different days under different positioning conditions, a rigid deformation was applied with the Elastix software for visualization purposes [33]. For these two mice, the DSCs, HDs and Δ COMs were calculated between the automatic and the manual segmentation at the first and last scan respectively. For the control mouse the DSCs were equal to 0.94 and 0.95, the HDs were equal to 0.4 mm and 0.3 mm, and the Δ COMs were equal to 0.3 mm and 0.2 mm respectively. For the TB mouse, the DSCs were equal to 0.95 and 0.94, respectively, the HDs 0.3 mm and 0.4 mm, and the Δ COMs were equal to 0.2 mm and 0.4 mm.

Muscle segmentation algorithm

The segmentations generated by the ‘forward’ two-step 3D U-Net based muscle segmentation algorithm were used to compare against muscle wet mass measurements. The first step of the 3D U-Net was trained in 5 hours and 31 minutes, the second step was trained in 6 hours and 44 minutes resulting in a total duration of 12 hours and

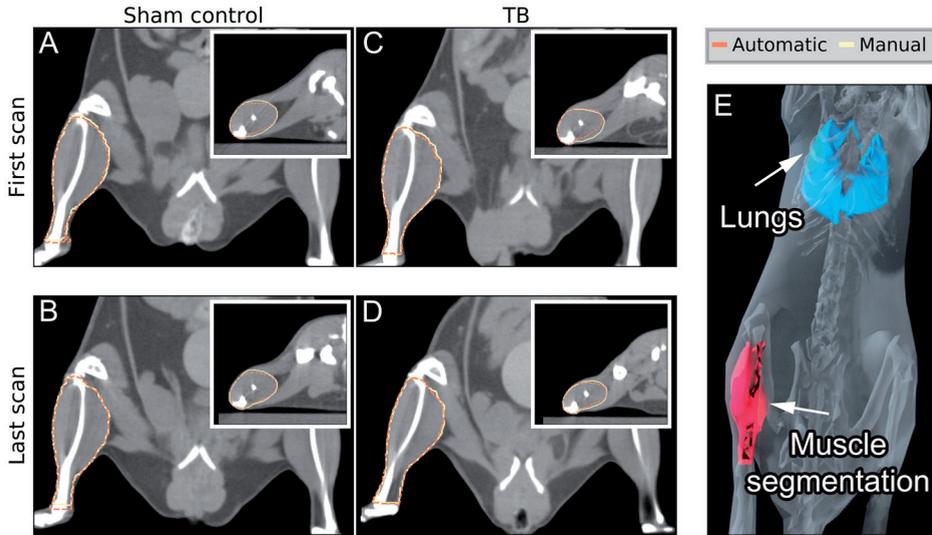


Figure 4. The automatic (orange) and manual (yellow) hind limb muscle segmentation on a μ CBCT slice for a sham control mouse and a tumor bearing (TB) mouse for the scan at the baseline measurement (upper row, A and C) and the last timepoint (bottom row, B and D). Axial μ CBCT slices are shown as subpanels in A-D. A three-dimensional model (E) of an example muscle segmentation is shown as illustration. In Supplementary Material, animations of all cross sectional CBCT slices are provided for both mice.

15 minutes. Performing the two-step 3D U-Net based automatic segmentation algorithm required an average calculation time of 6 seconds for one mouse, whereas the average manual segmentation time for one mouse was 17 minutes. Consequently, compared to manual segmentation, the machine-based automatic method reduces the time needed to segment hind limb skeletal muscle by more than a factor 150.

Discussion

In this chapter, we present a new method to non-invasively assess hind limb muscle mass in mice for cross-sectional and longitudinal purposes. It employs a deep learning algorithm trained to segment the lower muscle complex volume on μ CBCT images. This algorithm or model showed a high correlation with the actual muscle wet mass measurements. Furthermore, this deep learning application enables longitudinal evaluation of skeletal muscle mass changes in experimental mouse models, significantly reducing animal numbers and analytical workload.

Biomedical image segmentation methods are commonly trained on manually segmented datasets and validated against manual segmentations on a validation

dataset. After automatic segmentation, statements about the algorithm's performance are then based on specific parameters (e.g. DSC, HD, or ΔCOM) compared to the manual segmentation. Although we calculated these three quantitative parameters solely for a random sham control mouse and a random TB mouse at two time-points, we found a good agreement with the manual segmentations. The inter-observer variability and reproducibility of the manual segmentations evaluated in the image datasets of five mice showed similar results (DSC=0.95, HD=0.4 mm, ΔCOM =0.2 mm) compared to the quantitative metrics calculated between the manual and the automatic segmentation (DSC=0.95, HD=0.4 mm, and ΔCOM =0.4 mm).

In a more comprehensive quantitative analysis, we adopted a different approach to investigate whether a model can be trained on a dataset that was originally segmented by the algorithm. This means that the result of the 'forward' automatic segmentation algorithm was good enough to train a new deep learning model that in turn can segment μCBCT images with high accuracy. Using this reverse approach, we achieved mean high DSC scores of 0.93 and the HD and ΔCOM both had a sub-millimeter difference below 0.4 mm.

As experimental model validation, a good linear correlation ($R^2=0.92$) was found between the experimental wet muscle masses and the mass determined by our algorithm. The linear relationship was not equal to the unity curve which can be explained. First, not all soft tissues were extracted for the muscle wet mass measurements. Excision of the hind limb musculature was restricted to the gastrocnemius, soleus, plantaris, tibialis anterior and extensor digitorum longus muscles of both legs, as these can be collected intact and in a highly standardized manner. Consequently, remaining muscles such as the flexor digitorum longus, and the ankle plantar- and dorsiflexors, which contribute significantly to the lower hind limb muscle mass [34], contributed to muscle mass calculated by the algorithm but not the muscle wet weight measurements. However, in the training dataset and therefore also the automatic muscle segmentation, the whole muscle volume was segmented (in one leg) because the distinction between different muscle types requires enhanced image contrast which is not feasible with current μCBCT image quality. The muscle mass determined by our CT segmentation approach is therefore higher as it represents all, and not only the set of collected wet muscle masses. The high correlation between these two methods suggests that the relative differences between mice are maintained regardless whether total lower hind limb muscle mass or the cumulative wet mass of the standardized collected muscles are used.

Another source of uncertainty was the time between the last scan and the actual wet muscle mass measurements (day of sacrifice) which was ideally as short as possible (0 days), although this was not always possible for logistical reasons. The data extracted from the mice that were scanned at the day of sacrifice showed a clearly higher linear correlation ($R^2=0.96$) compared to the mice that were analyzed within

2 days after imaging ($R^2=0.80$). The latter may be attributed to continuous changes in muscle mass, including muscle wasting in the TB mice.

Previous studies, such as Ceelen *et al.* [20], already reported a good linear correlation between the CT-derived mass and the wet muscle mass for mice ($R^2=0.85$), but here the CT based muscle segmentation protocol was performed manually. The correlation in that study was comparable, but slightly lower than our linear fit based on the automatically determined CT-derived muscle masses ($R^2=0.92$). More recently, Pasetto and Olivari *et al.* [21] evaluated muscle atrophy in mouse models using μ CT imaging. However, no CT segmentation was applied in their work; instead muscle mass was evaluated differently according to 2D derived parameters including the perpendicular distance from the tibia half-length to the external hind limb muscle margin, and the distance from the upper extremity of the tibia to the medical malleolus.

The precision, accuracy, and reproducibility of DEXA in mice has already been investigated for in-vivo body parameters such as total bone mineral density, total body bone mineral, fat mass, and bone-free lean tissue mass [35, 36]. In Bunckinx *et al.*, DEXA was even proposed as the reference standard for measuring muscle mass in patients. Although DEXA provides longitudinal follow-up data of bone-free lean tissue mass with low radiation exposure, this technique cannot distinguish separate muscle groups such as the lower limb muscle complex, which is possible with the μ CBCT based algorithm. The use of μ CBCT as imaging modality in TB mice models, in contrast to DEXA, introduces additional longitudinal follow-up criteria from the same scan such as the tumor volume. MRI is an alternative high-resolution imaging technique, providing functional imaging with high soft tissue contrast. MRI does not require an imaging dose, in contrast to micro CT, which is an important consideration when radiation dose is a critical factor in the study setup. Although MRI provides a better soft tissue contrast than micro CT, the latter permits more rapid acquisitions at significantly lower cost. Additionally, it allows for tissue density quantifications, which makes micro CT a suitable imaging technique for longitudinal muscle mass assessment.

In the muscle wet mass procedure, only muscles were included that are excisable in a highly standardized way, whereas the manual and automatic segmentation methods determined the entire lower hind limb muscle complex volume. The muscle complex was segmented as a whole because the μ CBCT image reconstruction lacks image contrast to distinguish between individual muscle groups. In future studies, multi-modality image information, such as μ CBCT and MRI, could be combined with deformable image registration to investigate the segmentation feasibility of individual muscle groups, while the μ CBCT is required to convert image intensities to mass densities.

Lower limb muscle complex segmentation for multiple mice at different time points, risks at amounting into a time-consuming task in a preclinical study. To

illustrate this point: in our experiments, the manual segmentation by an experienced biologist of the lower limb complex on the μ CBCT images required on average 20 minutes for one mouse. The animals were scanned at 7 different timepoints on average, resulting in 581 μ CBCT scans, which would have taken more than 190 hours to segment using manual contouring. In contrast, the method described here required \sim 11 hours to segment the training dataset manually, around 12 hours to train the algorithm on GPU and only 6 seconds per automatic muscle segmentation. The latter would even enable real-time follow-up of muscle wasting during the in-vivo experiment.

Prior to implementation of the algorithm in other settings, additional validation steps may be required. In this study, the deep learning model was trained on a μ CBCT image dataset that was acquired with one specific imaging protocol at an X-ray tube potential of 50 kVp. However, further research is required to investigate whether the trained model is valid on imaging datasets that were acquired at different X-ray tube potential settings (e.g. with an additional scan at 90 kVp in dual-energy μ CBCT) or on imaging datasets that were acquired with different spectral filtrations [29]. Here, our training dataset consisted entirely of 129S2/SvPasCr1. As such, further investigations are required to evaluate how the trained model will segment the lower limb muscle in mice that have age-, strain, or pathological model-related differences in muscle size, compared to the mice that form the training dataset used in this study.

Nevertheless, we anticipate that the proposed algorithm is capable of segmenting muscle volumes in a variety of animals, as long as the study set-up (i.e. X-ray acquisition protocols and reconstruction settings) and animal characteristics are consistent over time, i.e. the period in which the training dataset is generated and applied to the non-segmented dataset. Therefore, we expect that for every study that uses a different imaging protocol or experimental animal species, the creation of a manually segmented training dataset and the re-training of the algorithm is required to achieve the best automatic segmentation outcome.

However, this algorithm, which can be re-trained depending on the study set-up, will facilitate the workflow tremendously because only the training dataset has to be segmented manually. Involving experimental models using μ CBCT or μ CT imaging in future work will be of interest to create a larger and more diverse training dataset of different animal species and settings scanned with different scan protocols and even different scanners to train a model that can be widely applied without the need for (extensive) re-training.

In conclusion, a non-invasive automatic algorithm was developed using artificial intelligence (two-step 3D U-Net) to segment skeletal muscle tissue in the lower limb complex of mice. The performance of the algorithm was in good agreement with the actual muscle wet mass measurements. This experimentally validated algorithm enables highly accurate non-invasive and automated longitudinal evaluation of skel-

etal muscle mass changes in mice with minimal operator involvement in the data analysis. It provides a unique possibility to collect large amounts of data from the mice, and to understand when and how muscle wasting starts and when it becomes life threatening. Furthermore, this method will allow a more accurate planning of experiments and reduce the number of mice needed for longitudinal experiments.

References

1. Janssen, I., et al., *Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr.* J Appl Physiol (1985), 1985. **89**(1): p. 81-8.
2. Evans, D.J., R. Murray, and A.H. Kissebah, *Relationship between skeletal muscle insulin resistance, insulin-mediated glucose disposal, and insulin binding. Effects of obesity and body fat topography.* J Clin Invest, 1984. **74**(4): p. 1515-25.
3. Periasamy, M., J.L. Herrera, and F.C.G. Reis, *Skeletal Muscle Thermogenesis and Its Role in Whole Body Energy Metabolism.* Diabetes Metab J, 2017. **41**(5): p. 327-336.
4. Schiaffino, S., et al., *Mechanisms regulating skeletal muscle growth and atrophy.* FEBS J, 2013. **280**(17): p. 4294-314.
5. Straub, R.H., et al., *Energy regulation and neuroendocrine-immune control in chronic inflammatory diseases.* J Intern Med, 2010. **267**(6): p. 543-60.
6. von Haehling, S. and S.D. Anker, *Prevalence, incidence and clinical impact of cachexia: facts and numbers-update 2014.* J Cachexia Sarcopenia Muscle, 2014. **5**(4): p. 261-3.
7. Anker, S.D., et al., *Wasting as independent risk factor for mortality in chronic heart failure.* Lancet, 1997. **349**(9058): p. 1050-3.
8. Mytelka, D.S., L. Li, and K. Benoit, *Post-diagnosis weight loss as a prognostic factor in non-small cell lung cancer.* J Cachexia Sarcopenia Muscle, 2018. **9**(1): p. 86-92.
9. Schols, A.M., et al., *Body composition and mortality in chronic obstructive pulmonary disease.* Am J Clin Nutr, 2005. **82**(1): p. 53-9.
10. Breiman, R.S., et al., *Volume determinations using computed tomography.* American Journal of Roentgenology, 1982. **138**: p. 329-333.
11. Kazemi-Bajestani, S.M., V.C. Mazurak, and V. Baracos, *Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes.* Semin Cell Dev Biol, 2016. **54**: p. 2-10.
12. Levine, J.A., et al., *Measuring leg muscle and fat mass in humans: comparison of CT and dual-energy X-ray absorptiometry.* J Appl Physiol (1985), 2000. **88**(2): p. 452-6.
13. Martin, L., et al., *Assessment of Computed Tomography (CT)-Defined Muscle and Adipose Tissue Features in Relation to Short-Term Outcomes After Elective Surgery for Colorectal Cancer: A Multi-center Approach.* Ann Surg Oncol, 2018. **25**(9): p. 2669-2680.
14. Prado, C.M., et al., *Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study.* Lancet Oncol, 2008. **9**(7): p. 629-35.
15. Shih, R., et al., *Lower limb skeletal muscle mass: development of dual-energy X-ray absorptiometry prediction model.* J Appl Physiol (1985), 2000. **89**(4): p. 1380-6.

16. Heier, C.R., et al., *Non-invasive MRI and spectroscopy of mdx mice reveal temporal changes in dystrophic muscle imaging and in energy deficits*. PLoS One, 2014. **9**(11): p. e112477.
17. Karlsson, A., et al., *Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI*. J Magn Reson Imaging, 2015. **41**(6): p. 1558-69.
18. Borga, M., et al., *Advanced body composition assessment: from body mass index to body composition profiling*. J Investig Med, 2018. **66**(5): p. 1-9.
19. Ceelen, J.J.M., et al., *Altered protein turnover signaling and myogenesis during impaired recovery of inflammation-induced muscle atrophy in emphysematous mice*. Sci Rep, 2018. **8**: 10761
20. Ceelen, J.J.M., et al., *Differential regulation of muscle protein turnover in response to emphysema and acute pulmonary inflammation*. Respir Res, 2017. **18**(1): p. 75.
21. Pasetto, L., et al., *Micro-computed tomography for non-invasive evaluation of muscle atrophy in mouse models of disease*. PLoS One, 2018. **13**(5): p. e0198089.
22. Park, J.S., et al., *Non-invasive tracking of disease progression in young dystrophic muscles using multi-parametric MRI at 14T*. PLoS One, 2018. **13**(10): p. e0206323.
23. Marcuzzo, S., et al., *Hind limb muscle atrophy precedes cerebral neuronal degeneration in G93A-SOD1 mouse model of amyotrophic lateral sclerosis: a longitudinal MRI study*. Exp Neurol, 2011. **231**(1): p. 30-7.
24. Baiker, M., et al., *Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data*. Med Image Anal, 2010. **14**(6): p. 723-37.
25. van der Heyden, B., et al., *Automatic multiatlas based organ at risk segmentation in mice*. Br J Radiol, 2019. **92**(1095): p. 20180364.
26. Gibbons, D.L., et al., *Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression*. Genes Dev, 2009. **23**(18): p. 2140-51.
27. Verhaegen, F., et al., *ESTRO ACROP: Technology for precision small animal radiotherapy research: Optimal use and challenges*. Radiother Oncol, 2018. **126**(3): p. 471-478.
28. Ma, C.M., et al., *AAPM protocol for 40-300 kV X-ray beam dosimetry in radiotherapy and radiobiology*. Med Phys, 2001. **28**(6): p. 868-93.
29. Vaniqui, A., et al., *The effect of different image reconstruction techniques on pre-clinical quantitative imaging and dual-energy CT*. Br J Radiol, 2019. **92**(1095): p. 20180447.
30. van Hoof, S.J., P.V. Granton, and F. Verhaegen, *Development and validation of a treatment planning system for small animal radiotherapy: SmART-Plan*. Radiother Oncol, 2013. **109**(3): p. 361-6.
31. Çiçek, Ö., et al., *3D U-Net: Learning dense volumetric segmentation from sparse Annotation*. arXiv, 2016.
32. Wang, C., et al., *A two-stage Unet framework for multi-class segmentation on full resolution image*. arXiv, 2018.
33. Klein, S., et al., *elastix: a toolbox for intensity-based medical image registration*. IEEE Trans Med Imaging, 2010. **29**(1): p. 196-205.
34. Charles, J.P., et al., *Musculoskeletal Geometry, Muscle Architecture and Functional Specialisations of the Mouse Hindlimb*. PLoS One, 2016. **11**(4): p. e0147669.
35. Halldorsdottir, S., et al., *Reproducibility and accuracy of body composition assessments in mice by dual energy X-ray absorptiometry and time domain nuclear magnetic resonance*. Int J Body Compos Res, 2009. **7**(4): p. 147-154.
36. Nagy, T.R. and A.L. Clair, *Precision and accuracy of dual-energy X-ray absorptiometry for determining in vivo body composition of mice*. Obes Res, 2000. **8**(5): p. 392-8.

10

Chapter 10

Summary, discussion and future perspectives

Summary and discussion

The overarching theme of this thesis is quantitative data extraction from CT images for radiotherapy research. I have chosen to focus on quantitative imaging, Monte Carlo simulation, organ-at-risk (OAR) segmentation, proton/photon radiation dose calculations for clinical radiotherapy (Ch. 2-7), and preclinical research (Ch. 8-9). Patient imaging is fundamental in a modern clinical radiotherapy workflow and the development of novel quantitative image data extraction methods are advantageous in multiple workflow stages; cancer screening, diagnosis, treatment planning, radiotherapy dose calculations, patient treatment positioning, and adaptive radiotherapy.

Clinical research

Research is desired to improve the CT image quality and the quantitative data extraction methods accordingly, however this is not always possible in clinical routine due to workflow limitations. In clinics, higher doses are required but are also a limiting factor for high quality CT, and policy or vendor restrictions are limiting research. CT simulation software provides a valuable alternative to mimic specific scanner characteristics and compare state-of-the-art analysis methods to the input reference of the simulation geometry. **Chapter 2** describes the development and validation of an open-access simulation tool ‘VOXSI’ in which the spectra and the detector response can be varied to find optimal scan parameters or settings depending on the final application. VOXSI is a user-friendly graphical user interface (**Figure 1**) that is able to perform SECT and DECT simulations and advanced post-processing: e.g. extract relative electron density (RED), effective atomic number (Zeff), and pseudo monoenergetic image (PMI) calculation [1-4]. Monte Carlo dose calculations require accurate material map information of the patient (i.e. density and material composition) which can be indirectly derived from the RED and the DECT-derived Zeff [5]. Here, simulation studies in digital anthropomorphic phantoms with input ground truth help to understand the behavior of DECT post-processing algorithms in more realistic geometries that are different from the basic tissue mimicking phantoms that lack anatomical structure.

During the development of VOXSI, a key goal was to make it available and simple for further extension by third parties in research collaborations. Currently, VOXSI is being extended by two national physics standard laboratories; for proton radiography (NPL, United Kingdom), and advanced noise modelling (PTB, Germany). Researchers currently use VOXSI to estimate Hounsfield Unit values of bone mimicking filaments for 3D printing and as a software toolkit to develop quality assurance (QA) phantoms for imaging. VOXSI has also been used as a CT simulation tool for imaging breathing motion artifacts in preclinical research [6]. It was the first research tool developed during my PhD research and served as a valuable tool for further studies.

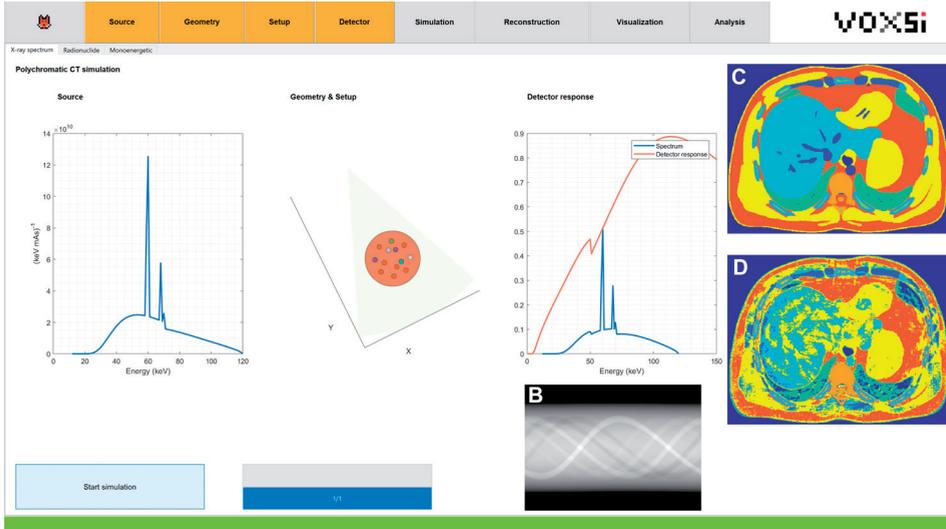


Figure 1. The VOXSI GUI interface - simulation window. Three panels (B-D) give an extra impression of possible simulation results; (B) the 120 kVp sinogram simulated from the RMI467 phantom, (C) the ground truth material map generated from a digital anthropomorphic phantom (liver), and (D) the segmented material map of the same digital phantom after DECT simulation and RED/Zeff post-processing.

While Monte Carlo dose calculations require specific material information, most clinical radiotherapy treatment planning systems require only densities or stopping power ratios (SPR) and apply a material conversion internally, if required. In order to do this, treatment planning systems typically request a single-energy (SE)CT to density conversion curve. These conversion curves are obtained via measurements in certified tissue mimicking phantoms such as the 062M phantom (CIRS Inc., Norfolk, VA) or the RMI 467 phantom (Gammex Inc., Middleton, WI). In these measurements, the tissue mimicking insert specifications provided by the phantom manufacturer are mapped to the measured HU values in the CT image reconstruction. However, this calibration curve depends not only on the X-ray tube potential (Figure 2B), but also on the patient size due to beam hardening (Figure 2A) which is often underrepresented in clinical routine.

In **Chapter 3**, the clinical validation of a new commercial CT image reconstruction algorithm (DirectDensityTM, Siemens Healthineers) was performed compared to the clinical protocol for megavoltage photon treatments in Maastricht, which depends on the object size and the X-ray tube potential (Figure 2A-B). The DirectDensity image reconstruction algorithm directly reconstructs the CT image in REDs independent from the X-ray tube potential. A simple measurement on the RMI 467 phantom, presented in Figure 2B, shows the X-ray tube potential dependency for two different

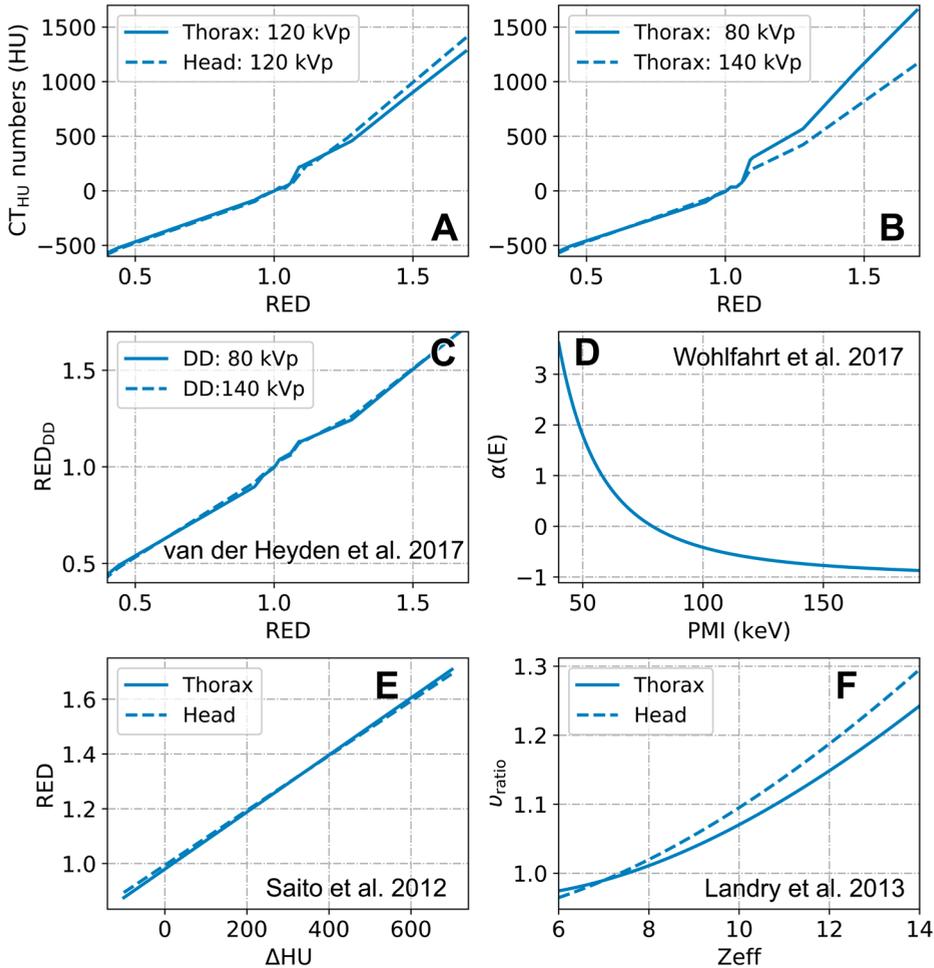


Figure 2. Conversion and calibration curves used in radiotherapy (research) to convert CT image intensities (either SECT or DECT) to material properties such as relative electron density (RED) or effective atomic number (Z_{eff}). (Abbreviations; (C) RED_{DD}, RED calculated by DirectDensity; (D) $\alpha(E)$: low- and high energy CT weighting factor in function of the pseudo-monoenergetic energy E, Supplement in Wohlfahrt et al. [7]; (F) ν_{ratio} : Equation 3 in Landry et al. [2]).

X-ray spectra (80 kVp and 140 kVp). When the same phantom acquisition is reconstructed with DirectDensity, the energy dependency is resolved (Figure 2C).

After commercial release, many questions were raised on how to use this algorithm in clinical routine because it has its advantages and disadvantages. The energy independency of the algorithm can, for example, ease the clinical implementation of automatic X-ray tube selection protocols for contrast-enhanced SECT scans in radiotherapy [8, 9] (CareKV, Siemens terminology) while eliminating the risk of wrong

SECT calibration curve selection in the TPS. Before the tumor and OAR doses of the planned treatment are evaluated via dose volume histogram (DVH) metrics, delineations on the CT are required to specify the organ boundaries and the planning margins. However, delineations on DirectDensity image reconstructions are not recommended for two reasons. First, the assumption made in the algorithm that every material exists of a water and bone-equivalent material combination is not valid for non-natural materials such as iodine contrast or metals (**Chapter 3**, Supplementary Material). Second, the algorithm applies a material composition in projection space to obtain sinograms of the effective water and effective bone thicknesses which are then combined to calculate a synthetic RED sinogram. The process to combine two separate sinograms could result in blurred sharp transitions between soft-tissue and bone in reconstructed image space. Both effects are detrimental for SECT image contrast, and therefore the 120 kVp DirectDensity image quality was qualitatively evaluated against the 120 kVp SECT in a short unpublished and internal study by three radiation oncologists. The radiation oncologists were asked to score the overall image quality, lesion and OAR contrast, image sharpness, and image artifacts according to the 5-point Likert scale [10]. Here, the 120 kVp SECT reconstruction scored unanimously better and approximately 10% higher than DirectDensity (Likert scale), agreeing with the reasoning behind the two technical limitations. This 10% difference was caused by lower scores on lesion and OAR contrast in the DirectDensity reconstruction. This means that an optimal clinical implementation of this algorithm requires two image reconstructions, one for delineation (120 kVp SECT) and one for MV photon dose calculations (DirectDensity). **Chapter 3** evaluated the MV photon dose differences between the conventional CT reconstruction and DirectDensity that were generally below 1% for the planning target volume. DirectDensity is also not recommended for particle therapy dose calculations in smaller body sites (e.g. head) because the algorithm development was optimized on abdominal sized phantoms which gave rise to a maximum RED overestimation of 6% for high-density cortical bone (**Chapter 4**). If DirectDensity is used for clinical segmentation tasks, studies are required to investigate the inter- and intra-observer variabilities in OAR or tumor delineation consistency between the conventional CT reconstructions (HU) and DirectDensity (RED/1000+1). Because the reconstruction units are different between both reconstructions, and to prevent window width/level outcome dependency, window leveling optimization must be considered.

Aside from the 120 kVp SECT to density conversion curve (**Figure 2A-B**) and the DirectDensity algorithm (**Figure 2C**), other commercial and non-commercial methods are also available to convert CT numbers to material characteristics. Several dual-energy CT (DECT) to material characteristic conversation methods are published [1, 11] or are available in commercial DECT post-processing packages such as *syngo.via* (Siemens Healthineers). In **Chapter 4**, the accuracy of several available

commercial and non-commercial SECT and DECT to density conversion methods was first investigated in a phantom study and then compared in four brain cancer patients. The tissue mimicking phantom study, including the 062M and the RMI467 phantoms, assessed differences in relative electron density between all conversion methods and compared the calculated values to the certified relative electron densities of the inserts, and for one of the patients, proton dose distributions were evaluated between the different conversion methods. The patient study reported dose differences compared to the clinical method (120 kVp, CT2RED conversion) because the relative electron density ground truth of the patient anatomy was lacking. The methods described by Saito et al. [1] and Hünemohr et al. [11] both required a calibration curve, but were the most accurate methods in our comparison which allows to speculate that these methods will provide the most accurate proton dose calculations in patients, where the ground truth is lacking.

A lacking material ground truth is always a weak point in comparison studies dealing with material composition. For proton therapy, research groups evaluate the CT-based material conversion accuracy in anthropomorphic phantoms by assigning measured SPRs relative to water [12, 13] or in heterogeneous tissue samples [14-16]. It is currently not possible to know the patient composition [17, 18] on the individual voxel level with sufficient accuracy to evaluate the performance of the various CT-based material property extraction methods. Phantoms for which the composition is well known do not faithfully represent real patient geometries and are also not ideal for evaluating the performance of these extraction methods in clinical situations. CT simulations of realistic digital patient volumes could help to tackle this problem (**Chapter 2**). Realistic digital humanoid phantoms such as the XCAT phantom that closely resembles an average person [19] can contribute substantially. Before the start of a CT simulation, the linear attenuation coefficient $\mu(E)$ must be specified in every voxel of the simulation geometry which is possible because all elemental compositions and tissue densities are known in advance. After simulation, the SECT or DECT post-processing (**Figure 2E** and **2F**) can be compared instantly to the ground truth which was initially desired as CT simulation input. Simulation allows the freedom to perform sensitivity studies with a variety of parameters (e.g. X-ray spectrum and filtration, detector response, image reconstruction, number of materials, and analysis methods) and to perform dose calculations that can be compared against a ground truth material map (simulation input). Also in pre-clinical radiotherapy research, realistic digital animal phantoms such as MOBY (mouse) [20] or ROBY (rat) [21] could help to define optimal imaging settings to maximize the image quality (i.e. contrast-to-noise, image-to-noise, or contrast agent visibility) or to optimize imaging spectra for more accurate dose calculations [22, 23].

In radiotherapy, the manual OAR delineation is a time consuming and error prone procedure prior to treatment planning. Efforts were already made in the liter-

ature to automatize this process with multi-atlas based image segmentation or Artificial Intelligence (AI) based image segmentation algorithms on conventional 120 kVp SECT scans, but with the emergence of new imaging modalities in radiotherapy such as DECT, clinics are not restricted to one specific image type anymore, so it was a logical next step to investigate whether different post-processing algorithms can improve the performance of state-of-the-art automatic segmentation algorithms. One of these post-processed datasets are pseudo-monoenergetic images (PMIs). The calculation of PMIs is a DECT-based post-processing technique wherein the two geometrically registered individual spectrum images are combined with an energy-dependent weighting factor $\alpha(E)$ (Figure 2D).

In Chapter 5, an AI-based algorithm for automatic OAR segmentation was qualitatively and quantitatively compared against a multi-atlas based algorithm on a range of PMIs ranging from 40 keV (Figure 3D) to 170 keV (Figure 3F). The results show that deep learning mostly outperforms multi-atlas based automatic OAR segmentation methods which agrees with both, the literature, and with the findings from the international AAPM thoracic auto-segmentation challenge, in which we competed [24, 25]. However, all previously published studies used conventional SECT image reconstructions as training or multi-atlas deformation dataset. One of the general findings was that the use of low energy PMIs for multi-atlas based image segmentation improved the OAR image segmentation accuracy and that the opposite was observed for deep learning based segmentation techniques. It would appear that the deformation algorithms, which are the ‘driving force’ behind the multi-atlas based algorithms, benefited from the contrast enhanced low energy PMIs while the increased noise levels at low energy perturbed the performance of the deep learning based segmentation algorithm. The PMIs did not require pre-processing and rescaling for the multi-atlas based algorithm but did require pre-processing for the deep learning based algorithm in the form of a HU normalization between two fixed HU thresholds. Since the HU scale is more stretched towards the lower pseudo monoenergetic energies, this could introduce a bias in the results. The multi-atlas based segmentation algorithm is less affected by fixed thresholds because the deformation algorithm allows the use of metrics that assume a relation between the probability distributions of the image intensities of the two datasets, e.g. the Mutual Information metric [26] in the Elastix software [27]. However, considering these limitations, the performance of the deep learning method is still excellent compared to other published studies and reports quantitative metrics close to the reported inter- and intra-observer variabilities [28-30].

The clinical automatic OAR segmentation work in this thesis is mainly applied on diagnostic CT image quality, but in image guided radiotherapy, CBCT images are also acquired prior to every treatment fraction. Here, OAR segmentation techniques are rather rare due to the inferior image quality compared to diagnostic CT and the

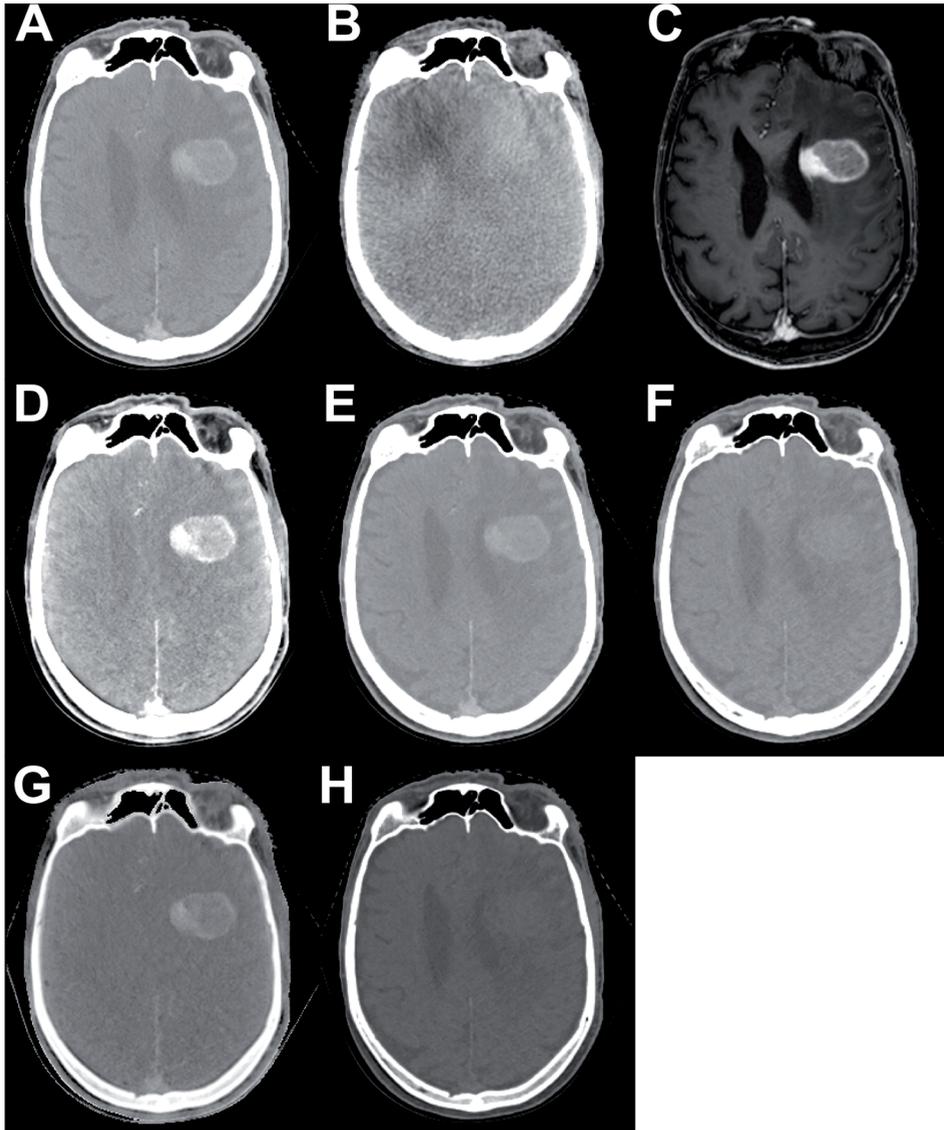


Figure 3. Images of several imaging modalities for a cancer patient that has a brain tumor. (A) SECT 120 kVp, (B) CBCT 125 kVp, (C) MR image T1 sequence, (D) PMI 40 keV, (E) PMI 79 keV, (F) PMI 170 keV, (G) Zeff, (H) RED. The Elastix software was used to register the images.

existence of deformation-based contour propagation algorithms. However, methods to enhance CBCT image quality are desired and should be investigated for many radiotherapy applications. In **Chapter 6**, the effect of the focal and off-focal radiation was quantified and in **Chapter 7**, an AI-based algorithm was trained to remove the scattered photon fluence from the X-ray projections.

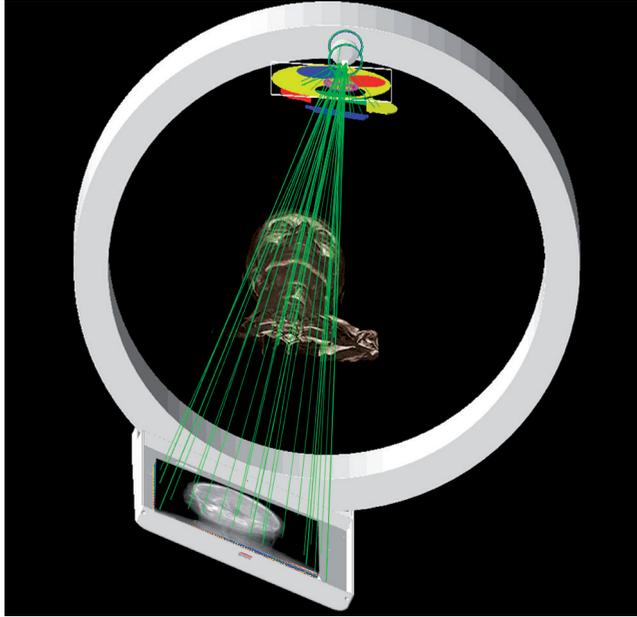


Figure 4. Sketch of the Monte Carlo simulation geometry of the medPhoton Imaging Ring on rails (IRr). Green lines depict photons produced in the X-ray tube.

A Monte Carlo simulation model of the medPhoton Imaging Ring system on rails (IRr) was developed in the Geant4 software (Figure 4) to investigate focal spot intensity distributions and, in practice, unmeasurable spectral and spatial characteristics of the off-focal spot radiation (Chapter 6). Image blurring at the level of the flat panel detector due to the finite focal spot size negatively influences the reconstructed CBCT image quality and therefore needs in-depth investigation [31, 32]. To simulate and evaluate the characteristics of off-focal spot radiation, this model included a refined X-ray tube simulation geometry including the interelectrode electric field between the cathode and the anode target. Primary intensity distributions of the emitted electron fluence from the cathode were estimated by iteratively solving an electron beamlet optimization problem to match simulations with focal spot slit camera measurements. This optimization problem was solved for the small and for the large focal spot and simulations were performed to investigate the characteristics of the off-focal spot radiation at diagnostic energies. To illustrate the versatility of advanced X-ray tube simulation models for future applications and studies, a small steel sphere was simulated at 120 kVp with three different source distributions (point source, small focal spot, and large focal spot) to investigate how the image quality was affected by blurring. The adjustable simulation model (kVp, target material, target mesh) described in Chapter 6 allows more insight in the focal and off-focal radiation and their effect on the image quality. This model could help to realistically

simulate focal spots, to assist the development of novel X-ray tubes, and to develop and evaluate the performance of projection-based deconvolution methods to improve image blurring artifacts [31, 32].

In addition to the focal spot size, photon scatter distributions majorly contribute to the inferior CBCT image quality compared to diagnostic fan-beam CT images [33]. Therefore, the Monte Carlo simulation model initially established in **Chapter 6** was extended in **Chapter 7** to separate primary and scattered photon fluences at the level of the flat panel detector to investigate scatter removal techniques. In earlier works, analytical and heuristic approaches were developed to correct for scattered photon fluences. The performance of these methods highly depends on the model parameterization and lacks robustness for non-uniform scatter distributions and variations in object position [34]. In **Chapter 7**, a deep variational autoencoder was trained on Monte Carlo simulation data of digital anthropomorphic head-and-neck phantoms simulated with non-isocentric imaging protocols at different virtual rotation centers. After model validation on simulated realistic digital patient phantoms, this study was the first to incorporate non-isocentric imaging trajectories, but more importantly this study was the first to apply a model trained on Monte Carlo simulation data on real patient acquisitions. The deep variational autoencoder correction method was compared against the current commercial implementation of the IRr and showed superior contrast-to-noise ratios, while taking the same computing time as the heuristic approach to correct the projections.

Preclinical research

In **Chapter 6** of the clinical research part of this thesis, a Monte Carlo model of an X-ray tube was developed to simulate more realistic focal spots intensity distributions and off-focal radiation. In image guided small animal precision irradiations, X-ray tubes are used to image (50-90 kVp) and to irradiate (225 kVp) the animals. Imaging is performed at lower energies with a small focal spot, and irradiations are performed in the orthovoltage energy range with a much larger focal spot ($\approx 2\text{mm}$). Because beam collimators are used during the irradiation with small exit windows (e.g. 1 mm), the collimator will act as a pinhole camera that projects the large focal spot intensity distribution on the animal. The complexity of precise dose calculations increases with smaller collimator exit windows. The model described in **Chapter 6** could be used to improve the accuracy of preclinical dose calculations for small beams, which illustrates the versatility of the simulation model [35-37].

Further preclinical research in this thesis focused on automatic μCBCT image segmentation. Although automatic image segmentation techniques were absent in preclinical work, they have the potential to facilitate and significantly speed up the preclinical workflow and data analysis for animal experiments. Compared to automatic image segmentation techniques for human radiotherapy (**Chapter 5**), very

little research was conducted in the preclinical field which limits the use and the general applications of these algorithms in preclinical studies. In **Chapter 8**, a multi-atlas based μ CBCT image segmentation algorithm was developed to segment OARs in mice. Compared with previously published studies that only provided organ contour estimations on mice data such as Baiker et al. [38] and Wang et al. [39], the overlap metrics (dice coefficient) for our automatic segmentation method almost doubled for the kidneys and the heart, 0.82 compared to 0.47, and 0.94 compared to 0.50, respectively.

In preclinical radiotherapy many steps have to be completed before the actual treatment delivery is performed, the animal fixation, μ CBCT imaging, manual tumor and OAR segmentation, and the dose calculations. The OAR segmentation is in this workflow one of the most time-consuming tasks for the human observer because the mice organs are relatively large compared to the submillimeter voxel sizes from the reconstructed μ CBCT images. Although it was not investigated in this thesis; automated planning or inverse planning would also improve the course of the experiment [40]. Both techniques, automatic OAR segmentation and automatic planning, will each lead to lower anesthesia times for the animal, reduce the risk of undercooling [40], and minimize the influence of a prolonged setup on the outcome of the experiment. Nevertheless, the algorithm presented in **Chapter 8** requires a large amount of computer power for the deformable image registrations to have reasonable OAR segmentation times. At the time this study was conducted, the algorithm segmented one mouse in 13 minutes on a CPU cluster of 144 cores. However, this average segmentation time was measured on a full-resolution mice dataset including several organs from different treatment sites (kidneys, heart, brain, and eyes). Preclinical treatments regularly treat only one body site (e.g. lung or brain) which means that the segmentation times can be reduced by cropping the dataset to the site of interest without information loss. The desktop computers that are available in preclinical research rooms usually lack high computing power which makes it challenging to implement the segmentation tool described in **Chapter 8** in preclinical routine. Currently, the research cluster of the Maastricht Physics research group is much larger (~800 cores) which would reduce the segmentation time significantly. However, adding more CPU computer power to improve segmentation speed does not solve the problem that limits preclinical implementation. Alternative fast algorithms, preferably running on a regular desktop computer, or maybe cloud computing solutions, are required to introduce the implementation of automatic OAR segmentation algorithms in preclinical radiotherapy software [41].

In **Chapter 9**, an automatic segmentation algorithm was developed to evaluate muscle atrophy (loss) and hypertrophy (growth) in μ CBCT images of cachexia developing mice based on AI software. Here, biologists from Maastricht University raised the question whether it was feasible to evaluate muscle loss or growth in a longitudi-

nal mice study with repetitive μ CBCT imaging, preferably in an automatic way. Because longitudinal studies require large animal cohorts being scanned at multiple time points and at regular time intervals, these longitudinal studies lead to a large number of datasets which all require manual intervention to evaluate e.g. the volume and mass of the lower limb muscle complex in this specific study. Often these studies also require large number of animals to be sacrificed to have sufficient measurement statistics at each time point, e.g. when cachexia is assessed by labor-intensive manual weighting of the excised muscles.

A deep learning based network (3D U-Net) was trained to segment the lower limb muscle complex in mice and was compared against muscle wet mass measurements, considered as reference method, after the last imaging timepoint. The automatic segmentation derived muscle mass showed a high linear correlation with the muscle wet mass measurements ($R^2=0.92$) and the automatic segmentation accuracy scored comparably to the manual segmentation accuracy, from both of which we concluded that the automatic muscle segmentation technique is a reproducible and an accurate method to perform longitudinal studies in mice. The development of this AI-based method was employed on a graphic processing unit making the automatic segmentation also approximately 150 times faster than the slow manual segmentation which required around 20 minutes per mouse. The μ CBCT dataset used to validate the automatic segmentation algorithm against the muscle wet mass measurements was part of a biological study that investigated different cell lines to develop an orthotopic lung tumor model in mice that induces cachexia. After finishing the model validation described in **Chapter 9**, the biological study was still on-going and in total, the automatic muscle segmentation algorithm was applied to 838 other mice μ CBCT datasets (December 2019), saving around 280 hours of contouring time by one human observer.

In summary, the biological study consisted of two phases. First, an orthotopic lung tumor animal model was developed that induced cachexia, and after the model development, the muscle mass changes were investigated in sham control and in tumor bearing mice cohorts. At the time of writing (Dec. 2019), biologists were optimizing the first phase of this biological study. The statistical sample size calculations to estimate the number of mice required in preclinical studies, considered a solitary lung tumor development rate of 70% (based on literature [42]) and a 30% drop-out due to mice reaching humane endpoints, 29% of the study cohort was assigned to the sham control group, and 71% were tumor bearing (TB) mice. During this first study phase, the automatic segmentation model was applied on datasets of 82 different mice. Methods to ease the use of longitudinal imaging of mice to investigate muscle mass changes will reduce the number of lab animals required to have sufficient model statistics.

Alternative methods that sacrifice sub-cohorts after specific timepoints require obviously much larger animal cohorts. Although less sacrificing timepoints (3 or 4) will be adopted in this approach, I estimate that only during the development phase of this specific biological study: (82 mice x 29% sham x 3 or 4 timepoints), between 82 and 109 less sham control mice were used, and (82 mice x 71% TB x 3 or 4 timepoints x 70% or 100% mice reach endpoint), between 140 and 267 less TB mice were used. Starting the second phase of the biological study, the method developed in **Chapter 9** will reduce the number of lab animals enormously, and worldwide, this method has to potential to reduce the use of lab animals with thousands every year.

Computed Tomography in Radiotherapy

Imaging is crucial in radiotherapy for OAR segmentation, tumor segmentation, prediction modelling, and radiotherapy dose calculations. In many clinics, non-optimal multi-purpose CT imaging protocols are commonly adopted. However, every task in radiotherapy has its own specific requirements. OAR segmentations are preferably performed on CT image with high soft-tissue contrast, potentially combined with MR imaging, whereas radiotherapy treatment planning requires quantitatively stable and accurate tissue information to achieve reliable dose calculations.

At the start of my PhD research, Maastricht just purchased a new dual-spiral DECT scanner (Somatom Confidence RT pro) which also invoked a market evaluation phase of the, at that time, unverified DirectDensity algorithm for MV photon dose calculations (**Chapter 3**). At that time (end 2016), most patients were treated with MV photons, as opposed to electron, proton, or orthovoltage photons, only one SECT scan was acquired per imaging slot and for some treatment sites multiple reconstructions were made with and without large-field-of-view reconstruction for dose calculation and delineation purposes, respectively. In IV iodine contrast-enhanced CT protocols, no additional CT scans were acquired prior to contrast injection because the dose calculation error in MV photon treatments is assumed to be clinically insignificant [43-45]. With the arrival of the DirectDensity image reconstruction algorithm combined with automated X-ray tube potential selection software (e.g. Siemens Care kV), CT scans were acquired in a range of energies (80-140 kVp) to optimize image contrast without the need of extra calibration curves per kVp in the TPS. Here, two image reconstructions are already desired, one for dose-calculations and one for volume delineation. As the DirectDensity algorithm is not recommended for volume delineations and because it was also limited to a small field of view of 50 cm instead of 65 cm, it was not suitable for all treatment sites which prevented the general clinical implementation at that time.

CT scans are not only used for structure delineation, but also for radiotherapy dose calculations. It was known from the literature that DECT-based methods outperformed SECT-based methods [7, 13] for proton therapy dose calculations, and

with the construction of the new proton therapy center in mind (Feb. 2019, Maastr Protonen Centrum, Maastricht), a more quantitative study was performed in **Chapter 4** to investigate several and SECT- and DECT-based conversion methods to derive tissue specific quantities needed for dose calculation. However, earlier studies mainly used commercial black-box implementations to generate e.g. SPR maps. In **Chapter 4**, non-commercial DECT-based methods were also explored, which allows the calibration process to be fully understood. These non-commercial DECT-based methods, such as Saito-15it [1] and Hünemohr-15it [11], outperformed conventional SECT based methods in the RED phantom study. Although the commercial DECT post-processing software (*syngo.via*) was available in Maastr, this software was not described in **Chapter 4** because a preliminary head-sized phantom study showed *syngo.via* RED errors up to 5% for high density bone inserts. These errors were caused because the commercial Hünemohr implementation in *syngo.via* lacked scanner, body site, and protocol specific calibration factors. In future releases of *syngo.via*, new features are desired to extend the DECT post-processing module and give clinics the opportunity to use their own fit calibration parameters, depending on the CT imaging protocol, to improve the density calculation accuracy and subsequently the dose calculations. These study results were extensively discussed with the development team of Siemens Healthineers.

DECT-based post-processing techniques such as PMI reconstructions (**Figure 2D-F**) provide a range of PMIs without the need of additional scanning. Radiological studies have demonstrated that low keV PMIs showed improved subjective image quality and improved lesion enhancement [46-49]. In clinical routine, I rather see PMIs as polished SECT scans with additional spatial-frequency filtering [50, 51] to improve contrast and image-to-noise ratios which could explain why in **Chapter 5** low energy PMIs did not provide better automatic OAR automatic contouring for the deep learning algorithm.

Wohlfahrt *et al.* found that the 79 keV PMI, also adopted in **Chapter 4**, is the optimal pseudo monoenergetic energy in dual-spiral DECT scanners (80/140 kVp) due to its superior image-to-noise ratio and its lower beam hardening [7]. At 79 keV the non-linear blending $\alpha(E)$ (**Figure 2D**) is approximately equal to zero, meaning that almost all data of the 140 kVp reconstruction is used (ignoring the spatial-frequency filtering [51]). Therefore, it seems logical that 79 keV obtained the lowest beam hardening for dual-spiral DECT scanners operating at 80/140 kVp. The 'optimal' energy was already determined for dual-spiral DECT scanners, but because $\alpha(E)$ depends on the effective linear attenuation coefficient of two basis materials (often water and cortical bone) at the low- and the high energy CT scan, further research is required to find optimal energy levels in dual-source DECT scanners or dual-layer DECT scanners. PMIs have the advantage that workflow-wise the calibration procedures are very similar to the conventional SECT protocols, which makes it easier to

implement a DECT-based method in clinical practice. Extra caution is needed in dual-spiral DECT as patient motion artifacts can introduce image artifacts due to incorrect deformations between the two scans (e.g. with bowel gas).

Treatment planning is often performed on a SECT image, sometimes combined with PET or MR imaging information. Additional information provided by DECT acquisitions makes the workflow (delineation and dose calculations) more complicated, but simultaneously provides more information which is beneficial for OAR and tumor segmentation [52]. A slider to change the PMI pseudo-monoenergetic energy (40-190 keV) in the TPS could potentially be used to optimize the image settings for delineation purposes and could even be an alternative for window levelling. Furthermore, DECT-based post processing techniques such as colored iodine maps superimposed on the CT dataset potentially improve the visibility and thus the segmentation of cancerous tissue [53-55]. DECT-based RED and Zeff maps can be used to perform better tissue identification and improve dose calculations, more specifically for treatments [56]. This discussion also implies to preclinical radiotherapy research [57, 58].

Artificial Intelligence in radiotherapy

“Artificial Intelligence (AI)” is since 2017 one of the most commonly used buzzwords in radiotherapy research. At the start of my PhD research mid 2016, I started the development of an OAR segmentation algorithm for CT datasets. After a literature survey I found that multi-atlas-based image segmentation methods were the best algorithms at that time [59-63]. AI was considered as a very complex and unreachable piece of software that was indirectly linked to Google-sized companies and specialized computer scientists. In 2014, a Vision 20/20 paper about automated image segmentation technique for radiotherapy was not even mentioning AI [30] and the participants in automatic contouring challenges only presented results based on their own multi-atlas-based image segmentation methods [64]. A quick survey on PubMed with search terms as “deep learning” and/or “artificial intelligence” in the title or abstract learned me that from 2000 to 2014, 1217 publications were found and from 2015 to 2019 (until Dec.), 10372 publications were found, almost ninefold in a period that was three times shorter. Only since the release of TensorFlow in November 2015 and the emergence of affordable graphical processing unit cards, AI suddenly became more accessible to the wider scientific community, in medicine and thus radiotherapy. In the following automatic segmentation challenge in 2017, where I also participated, the algorithms were already predominantly AI-based and outperforming the multi-atlas-based image segmentation methods [25]. After this challenge, it was clear that AI-based methods instead of multi-atlas-based image

segmentation methods would be the focus of my studies moving forward. However, the efforts in multi-atlas based image segmentation methods and thus deformable image registration techniques are not useless in the wake of AI. In radiotherapy, atlas based image segmentation techniques are still applicable for tumor and OAR contour propagation from planning CT to daily CBCT, between the daily CBCTs [65-69], or to propagate contours between different phases of a 4D CT scan [70, 71]. For these applications, the atlas-based method performs well because the manual segmentations, delineated on an earlier dataset from the same patient, only show small changes during radiotherapy course.

In the clinical research part of this thesis, AI was used for automatic organ-at-risk segmentation in dual-energy CT derived images (**Chapter 5**), and an AI implementation was developed to remove scattered photons from X-ray projections in non-isocentric CBCT acquisitions. A growing body of literature suggests the large number of possibilities that AI-based applications can bring to medical physics and radiotherapy [72]. In a relatively short period of 3 years, AI has been used for single- and multi-modality image segmentation, synthetic CT image generation from MR [73], image denoising or noise reduction in low dose CT [74, 75], image reconstruction [76], material mapping in dual-energy CT [77], focal spot deconvolution [31], projection-based X-ray scatter removal [78-81], or markerless tracking stereotactic body treatments [82].

Much work on the potential of AI in radiotherapy has been carried out, however, a scientifically critical view on AI is required. Solutions must be thought out carefully and AI may only solve the problem in hand partially. This wave of AI applications, which is seen as part of the fourth industrial revolution, should be met with as much scientific rigor as excitement. It appears to be advantageous and appropriate to replace humans in repetitive, error-prone, and monotonous time-consuming radiotherapy tasks by computers, but currently, AI can only interpret data on which the algorithm is trained and at the time of writing is still lacking true or general intelligence. I experienced this during our automatic OAR segmentation study in **Chapter 5**. Here, our AI algorithm was trained and tested on a small cohort of adult patient data but failed completely when it was applied on pediatric case. No or wrong segmentation results were found by the AI algorithm while the multi-atlas-based algorithm still did a good automatic OAR contour segmentation due to the deformation algorithms. This personal experience serves as a good illustrative example that shows how careful we must handle AI algorithms in healthcare. How well will the AI algorithm perform when it faces a problem it hasn't seen in the training dataset? It is currently not possible to test AI algorithms with every possible input so it is not yet possible to rely on AI algorithms for anything other than finely defined single tasks that a human can still correct. For the foreseeable future, a final check by a human observer will still be required when AI is applied in a clinical radiotherapy workflow.

The type of algorithms give rise to many other open questions; can AI algorithms handle low-energy CT data (e.g. 80 kVp) when it was trained on conventional 120 kVp CT data; can AI algorithms be trained on CT data of manufacturer 'A' be applied on CT data of manufacturer 'B'; what about all different image reconstruction and correction algorithms, etc.? Testing can increase our confidence in these methods but can and should not tell us to stop being critical of their ability. Although researchers aiming to solve a difficult problem use the term 'big data' too frequently and flippantly, here, multi-institutional and cross-country derived big datasets have the potential to tackle this problem. Returning to the illustrative example of the pediatric OAR segmentation in **Chapter 5**, although 3D data was used, the number of patients used in the training dataset was too small to have a generic working algorithm. When the training dataset was larger and included pediatric data, for example, this allows for speculation that the algorithm would have achieved a higher segmentation accuracy for this pediatric patient as well.

An obvious question at conferences following a presentation on automatic segmentation is "What should happen with all those unemployed medical doctors and radiotherapy technologists due to AI?". As with automation in any technical field, there are more useful and less repetitive tasks that require experienced human attention in the clinic. For example, when the radiation oncologist has more time available due to AI, the quality of care will improve and would e.g. benefit from personalized medicine or treatment decision support tools [83-85]. Here, the radiation oncologist introduces possible treatment techniques or schemes before the start of the radiotherapy treatment. The patient can evaluate the techniques or treatment schemes at home and come back to discuss the final decision together with the radiation oncologist who will spend more time in consultations. Since proton therapy became clinically available in MAASTO last year and because the Netherlands adopts a model-based approach [86], whereby the dose distributions of photon and proton treatment must be compared to see which treatment would be better suited for the patient, this doubles the work for radiation oncologists, but is better for the patient. Automated OAR segmentation algorithms can already eliminate the OAR segmentation burden in the planning stage to speed up the workflow, and free more time to optimize the treatment plan quality together in collaboration with other clinics.

AI should definitely be considered as a technique to help with clinical repetitive and time-consuming tasks in radiotherapy. AI may not perform better than the data it's been trained on, but it could help the rollout standardized segmentation guidelines in all radiotherapy clinics, also in less developed countries, and change the way medicine is performed.

Artificial intelligence in preclinical studies

Following the work performed during my MSc., there was an opportunity to automate the time-consuming image segmentation task in the preclinical research field as no techniques were available. **Chapter 8** was the first study to apply more advanced automatic segmentation algorithms on μ CBCT images, such as multi-atlas based image segmentation, on mice OARs. Next, Chapter 9 applied AI techniques to segment volumes of interest on μ CBCT images in a longitudinal study. Both chapters showed that automated image segmentation techniques do not only facilitate the workflow during the experiment, it also improves the wellbeing of the animal due to lower anesthesia times, significantly less animals needed to be used in longitudinal studies (according to the 3R paradigm, 'Reduce, Refine, Replace'), and finally it also reduces the time consuming manual segmentation burden of the biologist or lab technician after the imaging was performed. Most large-scale research institutions or universities already have in-house knowledge to start the development of algorithms to facilitate the preclinical workflow. Interdisciplinary collaborations between departments, e.g. **Chapter 9**, can lead to interesting findings, new research ideas, and useful applications for all parties.

Limitations and considerations

Planning CT

Today various planning CT modalities and post-processing algorithms are available to perform SECT or DECT imaging for radiotherapy depending on the treatment site, the required imaging FOV, and the irradiation treatment (kV photons, MV photons, protons). Imaging protocols and the use of CT images in radiotherapy workflows should be well thought out, depending on the irradiation requirements. Especially for low energy photon and proton radiotherapy, where many studies already showed the improved dose calculations accuracy for DECT [7, 13, 16, 87-89].

Originally, DECT was a dedicated technique for radiology applications for improved contrast imaging [90] used for lung perfusion imaging [91], the diagnosis of cardiovascular diseases [92], the detection of calcified plaques, or the detection of renal stones [93]. All these DECT applications require much smaller FOVs than the FOVs desired for radiotherapy applications, which require full patient anatomy for dose calculations. This limited FOV of dual-source DECT systems limits the clinical implementation of DECT for specific treatment sites and the technique must become more mature e.g. the use of DECT images in the TPS. Alternatively, sub-optimal dual-spiral or twin-beam DECT approaches could be used [94]. Commercial treatment planning systems currently do not provide options to process more-advanced

DECT image data for structure delineation and dose calculations. Only refined SECT-mimicking images such as PMIs can currently be used in commercial TPSs to perform dose calculations according to the conventional CT calibration protocols. Combined efforts of the imaging and the treatment device manufacturers are desired to have commercial DECT-friendly systems to boost the clinical usage of more advanced DECT technology in radiotherapy.

Organ-at-risk (OAR) segmentation

This thesis discussed automatic image segmentation techniques for preclinical μ CBCT images (**Chapter 8** and **9**) and for clinical planning CT images (**Chapter 5**). Automatic segmentation results require careful interpretation as training, testing, and validation is often performed on data from the same institution, acquired with the same imaging protocols, contouring guidelines, etc.

Fixed HU thresholds are required in the data pre-processing step to convert CT images (HU) to normalized data (0-1) that can be processed by the AI algorithms. These fixed thresholds, based on energy dependent HUs, are a weak point of AI-based segmentation methods which could limit the performance of the automatic segmentation algorithm. Further training and validation of these algorithms on more diverse, and preferably, external datasets are desired to evaluate the general performance. Therefore, I encourage the organization of automatic segmentation challenges with independent outcome validation [25]. These challenges are actually one of the few solutions to compare the performance of several segmentation algorithms on the same external institutional dataset [25].

With the introduction of more advanced imaging modalities such as DECT, it will start to become more difficult to standardize image data. E.g. CT images acquired at different tube potentials will have a different Hounsfield Unit scaling which could reduce the performance of AI algorithms. I would encourage researchers in future studies to take one step back and improve the data normalization and standardization before showing algorithm performance. In automatic CT segmentation research, (rescaled) energy dependent HUs are still used as input for the AI algorithms, even if there are already existing software solutions and calibration procedures to convert HUs to material specific REDs with DirectDensity, *syngo.via* RED, or with a simple CT2RED conversion curve. This way, standardization could be performed on physically extracted data that is independent from the CT modality (SECT or DECT) or the X-ray tube potential. Further research is required as important information could get lost following a RED-based standardization approach. E.g. the reduced appearance of the iodine contrast comparing conventional 120 kVp SECT with DECT-based *syngo.via* RED (**Figure 3A** and **Figure 3H**) could reduce the tumor segmentation performance in future applications.

In the literature, the performance of OAR segmentation algorithms is mainly evaluated based on overlap metrics such as the volumetric Dice Similarity Coefficients (DSC), or distance metrics such as the Hausdorff Distance (HD) or the center of mass displacement (ΔCOM). These metrics give only an idea about the automatic image segmentation quality but do not always represent clinical relevance [95]. The opinion of medical doctors is required to evaluate the clinical significance and contour acceptability. Metrics that correlate with delineation time should be investigated and reported together with overlap metrics and distance metrics. The surface DCS [96] and the added contour path length [97] are two of these recently introduced metrics that show higher correlation with manual contour adaptation time compared to the conventional metrics.

Cone-Beam CT

This thesis devoted two chapters to the Monte Carlo simulation of a novel CBCT scanner that is installed as imaging modality in Maastricht Protonen Centrum. This ImagingRing™ system on rails (IRr) (medPhoton GmbH, Salzburg, Austria) is able to perform non-isocentric image acquisitions with an independently rotatable and collumatable X-ray source and flat panel detector, and was simulated to create a training database of primary and scattered photon transmission projections through realistic digital anthropomorphic patient geometries. This approach required Monte Carlo simulations as the simultaneous measurement of primary and scattered photons is impossible over such a large area with current detector technology. The simulated database was then used for training of a variational autoencoder to correct for scatter in real patient acquisitions. The X-ray tube geometry was also modelled to simulate focal and off-focal radiation effects. This task, which at first sight was a rather academic assignment, finds several useful use cases in imaging research and even preclinical precision irradiations at orthovoltage energies. During this simulation work, only slit measurements of the horizontal and vertical focal spot profiles were used in order to simulate more realistic and irregularly shaped focal spot intensity distributions. Further model validation is lacking and needs to be performed in future studies, e.g. half value layer measurements and anthropomorphic phantom simulation with known material compositions.

In the future, similar to the AI-based scatter correction algorithm, a database could be created with simulations to train a convolutional neural network to deconvolve realistic focal spot distributions without the use of point-spread functions [31]. Also the IRr system manufacturer aims to perform dual-energy CT imaging by fast kV switching while simultaneously rotating the spectral filtration wheel consisting of three different spectral filters and air. Although it was not investigated in this thesis, the IRr Monte Carlo simulation model was extended in 4D to simulate these very

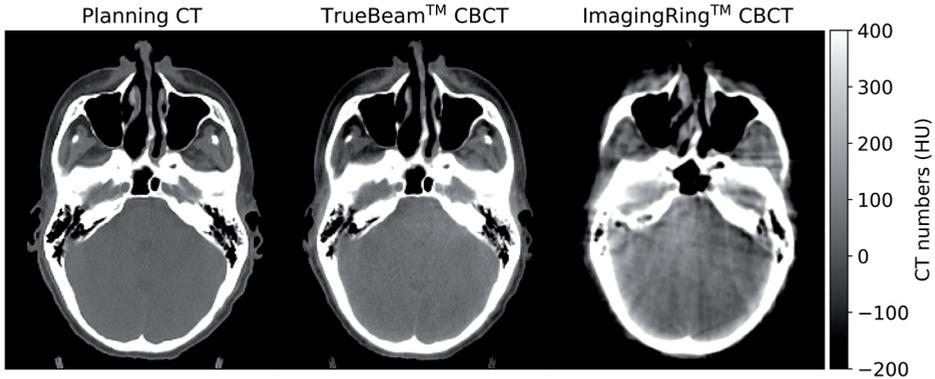


Figure 5. Image comparison between 120 kVp diagnostic planning CT, 125 kVp TrueBeam™ CBCT, and the 120 kVp ImagingRing™ CBCT.

specific imaging trajectories per projection, optionally with a 4D digital anthropomorphic phantom (e.g. 4D XCAT phantom [19]) to anticipate on further research.

In the Maastricht Protonen Centrum, IRr imaging is only performed for patient alignment. Efforts to improve image quality are needed to enable applications such as online proton dose recalculations. However, at this stage, the general reconstructed IRr images are full of image artifacts, sensitive to patient motion, and produces inconsistent phantom images of tissue mimicking phantoms such as the RMI 467 phantom (Gammex) and the 062M phantom (CIRS) that are required for CT to density conversions. The CBCT image quality that is obtained with the IRr system is currently inferior to linac mounted CBCT image quality and diagnostic fan-beam CT image quality. However, the IRr system allows for non-isocentric image acquisitions and should theoretically be able to focus on one treatment site by collimating the beam which reduces the CBCT imaging dose (not a clinical product) which is of interest in fractionated radiotherapy and in line with the ALARA principle [98-100].

Considering the relatively large proton dose differences and proton ranges that are already found between diagnostic SECT and DECT imaging [56], the dose differences between diagnostic CT and IRr will exceed far beyond clinical acceptability. For plan adaptation and online proton dose recalculations, in-room CT or even in-room DECT on rails should currently be considered as alternative for CBCT [101]. In-room CT with robotic couch for proton therapy should enable more accurate and faster imaging, is less subject to patient motion, and will even speed-up the imaging procedure because current CBCT solutions are relatively slow. Moving the ImagingRing™ from its rest position to its starting imaging position takes around one minute and the image acquisition itself another minute.

Future perspectives

In the near future, I predict radiotherapy clinics to continue slowly changing their CT scanners and standard imaging protocols towards DECT imaging. Before DECT is more widely available, image reconstruction algorithms that convert SECT scans independently from the X-ray tube potential to densities (e.g. DirectDensity) are a good alternative as this allows for optimal patient specific X-ray tube potential selection protocols to optimize image contrast. However, DirectDensity only converts to RED in a relatively small 50 cm field of view, so efforts from commercial partners are desired to extend to field of view for larger body sites (e.g. head-and-neck, pelvis, breast) and to perform algorithm adaptations to convert CT images into mass density images which are desired for e.g. AcurosTM XB dose calculations in Eclipse TPS (Varian Medical Systems). Some radiotherapy specific challenges in DECT imaging need to be resolved before more clinics will consider the purchase of a more expensive DECT scanner. In dual-source DECT scanners, the limited field-of-view of 33 cm limits the usage of DECT for lung, breast, pelvic, and head-and-neck cancer patients. Software or hardware solutions (e.g. large bore) should be investigated to provide extended field of view imaging with consistent image intensity values. The limited field-of-view significantly restricts external beam therapy applications, but 33 cm is sufficient for intra-operative therapies such as brachytherapy [87]. Currently, DECT images cannot be processed separately by the treatment planning system and concerted efforts from the imaging and TPS manufactures are required to incorporate post-processed DECT data in the treatment planning system to improve OAR and tumor visibility for delineation, and to improve the accuracy of the radiotherapy dose calculations. Some AI algorithms will in the short term become a standard technology in radiotherapy and will reduce the time-consuming repetitive tasks such as OAR segmentation, but other AI-based applications still need further investigation. To boost automatic segmentation performance of OARs and even tumors, research is required to incorporate multi-modality information in the AI algorithms. CT and MRI datasets could be combined to improve automatic OAR segmentation as some OARs are barely visible on CT (e.g. hippocampi, optic chiasm, hypothalamus). In a recent European Particle Therapy Network (EPTN) consensus, a consensus-atlas for CT- and MR based contouring in neuro-oncology was provided [102] and recommendations were made to standardize OAR segmentation to improve the future development of normal tissue complication probability (NTCP) models that strongly depend on uniform delineation. AI methods that combine multimodality image data could incorporate this consensus atlas to better standardize neuro-oncology. Contrast-enhanced DECT and MRI data could be combined for automatic tumor segmentation, potentially in combination with functional imaging data such as PET. Other AI-based algorithms that aim to improve the CT image reconstruction

quality with denoising algorithms or scatter removal algorithms have the potential to be implemented in a clinical workflow when they are trained and extensively validated on sufficient large datasets to incorporate less common clinical user cases.

I also predict that technical advances such as automatic image segmentation for preclinical research will increase in the near future. In recent years, neural network architectures have been investigated and validated extensively to replace human tasks and improve existing methods. Because these network architectures already exist, they can be applied to preclinical research with the aid of collaborations between radiotherapy specialists and preclinical researchers. Requirements for preclinical research differ from clinical OAR segmentation (e.g. segmentation of the lower hind limb muscle complex in mice (**Chapter 9**) to correlate with outcome prediction and cancer progression), which makes it ambitious to train generally applicable models. However, in contrast to radiotherapy segmentation tasks, preclinical research is performed on cohorts that consist of very similar subjects (age, strain, weight, living conditions, medical prehistory), so a relatively small number of training datasets should be sufficient to prevent time consuming manual segmentation by human observers.

Research is still ongoing to improve the hardware technology of CT or CBCT imaging systems. In my opinion, correction algorithms are technology and platform dependent, and therefore should be only pursued with current technology in mind. If a new technology is developed that avoids the need for a specific correction, then the new technology should be embraced, and its limitations investigated. AI algorithms to artificially improve CBCT image data are maybe not desired in clinical routine where submillimeter accuracy is aimed for, as these correction algorithms are only predicting how a patient anatomy looks like without a reference. Although AI correction algorithms can generate highly accurate-looking data (e.g. image space mapping from CBCT to diagnostic CT [103]), people need to know that they make intermediate decisions on data whose accuracy and performance cannot directly be evaluated by a human observer. This is different for AI algorithms used for automatic OAR segmentation of which the performance is and must be evaluated compared to the anatomical knowledge of an experienced human observer. Here, the AI results are also overlaid on the original CT reconstruction that medical doctors use to make clinical decisions.

To improve current CT and CBCT imaging systems, further hardware developments in the energy-resolving photon-counting detector technology would offer a robust solution to perform multi-energy CT imaging. This technology potentially allows for multi-energy CT imaging in one acquisition with only one photon counting detector array and one X-ray tube, which would directly eliminate the motion artifacts between two acquisitions in dual-spiral DECT, and would eliminate the field of view limitation in dual-source DECT, while providing multi-energy data.

Photon-counting detector technology commercialized by the MediPix3 collaboration [104-106], or the currently ongoing MediPix4 collaboration, have an enormous potential to replace current detector technology in CT and in CBCT modalities because faster readout times are obtained, less particles are needed for imaging, and devoted hardware corrections can be employed eliminating scatter (e.g. incidence angle and time of flight corrections).

References

1. Saito, M., *Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship*. Med Phys, 2012. **39**(4): p. 2021-30.
2. Landry, G., et al., *Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients*. Phys Med Biol, 2013. **58**(19): p. 6851-66.
3. Saito, M. and S. Sagara, *A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body*. Med Phys, 2017. **44**(6): p. 2293-2303.
4. Yu, L., et al., *Virtual monochromatic imaging in dual-source dual-energy CT: radiation dose and image quality*. Med Phys, 2011. **38**(12): p. 6371-9.
5. Landry, G., et al., *Simulation study on potential accuracy gains from dual energy CT tissue segmentation for low-energy brachytherapy Monte Carlo dose calculations*. Phys Med Biol, 2011. **56**(19): p. 6257-78.
6. Vaniqui, A., et al., *On the determination of planning target margins due to motion for mice lung tumours using a four-dimensional MOBY phantom*. Br J Radiol, 2019. **92**(1095): p. 20180445.
7. Wohlfahrt, P., et al., *Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans*. Int J Radiat Oncol Biol Phys, 2017. **97**(2): p. 427-434.
8. Kalender, W.A., et al., *Application- and patient size-dependent optimization of X-ray spectra for CT*. Med Phys, 2009. **36**(3): p. 993-1007.
9. MacDougall, R.D., P.L. Kleinman, and M.J. Callahan, *Size-based protocol optimization using automatic tube current modulation and automatic kV selection in computed tomography*. J Appl Clin Med Phys, 2016. **17**(1): p. 328-341.
10. Likert, R., *A technique for the measurement of attitudes*. Archives of Psychology, 1932. **140**: p. 1-55.
11. Hunemohr, N., et al., *Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates*. Phys Med Biol, 2014. **59**(1): p. 83-96.
12. Grant, R.L., et al., *Relative stopping power measurements to aid in the design of anthropomorphic phantoms for proton radiotherapy*. J Appl Clin Med Phys, 2014. **15**(2): p. 4523.
13. Wohlfahrt, P., et al., *Evaluation of Stopping-Power Prediction by Dual- and Single-Energy Computed Tomography in an Anthropomorphic Ground-Truth Phantom*. Int J Radiat Oncol Biol Phys, 2018. **100**(1): p. 244-253.
14. Bar, E., et al., *Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples*. Med Phys, 2018. **45**(1): p. 48-59.
15. Mohler, C., et al., *Experimental verification of stopping-power prediction from single- and dual-energy computed*

- tomography in biological tissues*. Phys Med Biol, 2018. **63**(2): p. 025001.
16. Taasti, V.T., et al., *Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples*. Phys Med Biol, 2017. **63**(1): p. 015012.
 17. Woodard, H.Q. and D.R. White, *The composition of body tissues*. Br J Radiol, 1986. **59**(708): p. 1209-18.
 18. White, D.R., H.Q. Woodard, and S.M. Hammond, *Average soft-tissue and bone models for use in radiation dosimetry*. Br J Radiol, 1987. **60**(717): p. 907-13.
 19. Segars, W.P., et al., *4D XCAT phantom for multimodality imaging research*. Med Phys, 2010. **37**(9): p. 4902-15.
 20. Segars, W.P., et al., *Development of a 4-D digital mouse phantom for molecular imaging research*. Mol Imaging Biol, 2004. **6**(3): p. 149-59.
 21. Tsui, B.M. and W.P. Segars. *C04404: 4D Digital Rat Whole Body (ROBY) Phantom*. 2004; Available from: <http://jhu.technologypublisher.com/technology/16325>.
 22. Schyns, L.E., et al., *Optimizing dual energy cone beam CT protocols for pre-clinical imaging and radiation research*. Br J Radiol, 2017. **90**(1069): p. 20160480.
 23. Almeida, I.P., et al., *Exploring the feasibility of a clinical proton beam with an adaptive aperture for pre-clinical research*. Br J Radiol, 2019. **92**(1095): p. 20180446.
 24. Lustberg, T., et al., *Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer*. Radiother Oncol, 2018. **126**(2): p. 312-317.
 25. Yang, J., et al., *Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017*. Med Phys, 2018. **45**(10): p. 4568-4581.
 26. Thevenaz, P. and M. Unser, *Optimization of mutual information for multiresolution image registration*. IEEE Trans Image Process, 2000. **9**(12): p. 2083-99.
 27. Klein, S., et al., *elastix: a toolbox for intensity-based medical image registration*. IEEE Trans Med Imaging, 2010. **29**(1): p. 196-205.
 28. Faggiano, E., et al., *An automatic contour propagation method to follow parotid gland deformation during head-and-neck cancer tomotherapy*. Phys Med Biol, 2011. **56**(3): p. 775-91.
 29. Deeley, M.A., et al., *Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study*. Phys Med Biol, 2011. **56**(14): p. 4557-77.
 30. Sharp, G., et al., *Vision 20/20: perspectives on automated image segmentation for radiotherapy*. Med Phys, 2014. **41**(5): p. 050902.
 31. Kuntz, J., et al., *Focal spot deconvolution using convolutional neural networks*. Proc SPIE Medical Imaging 2019: Physics of Medical Imaging, 2019. **10948**.
 32. Nagesh, S.V.S., et al., *Focal Spot Deblurring for High Resolution Direct Conversion X-ray Detectors*. Proc SPIE Int Soc Opt Eng, 2016. **9783**.
 33. Siewerdsen, J.H. and D.A. Jaffray, *Cone-beam computed tomography with a flat-panel imager: magnitude and effects of X-ray scatter*. Med Phys, 2001. **28**(2): p. 220-31.
 34. Suri, R.E., et al., *Comparison of scatter correction methods for CBCT*, in *Medical Imaging 2006: Physics of Medical Imaging*. 2006.
 35. van Hoof, S.J., J.B. Verde, and F. Verhaegen, *Dose painting by dynamic irradiation delivery on an image-guided small animal radiotherapy platform*. Br J Radiol, 2019. **92**(1095): p. 20180744.
 36. Granton, P.V. and F. Verhaegen, *On the use of an analytic source model for dose calculations in precision image-guided small animal radiotherapy*. Phys Med Biol, 2013. **58**(10): p. 3377-95.
 37. Verhaegen, F., et al., *A review of treatment planning for precision image-*

- guided photon beam pre-clinical animal radiation studies.* Z Med Phys, 2014. **24**(4): p. 323-34.
38. Baiker, M., et al., *Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data.* Med Image Anal, 2010. **14**(6): p. 723-37.
 39. Wang, H., D.B. Stout, and A.F. Chatziioannou, *Estimation of mouse organ locations through registration of a statistical mouse atlas with micro-CT images.* IEEE Trans Med Imaging, 2012. **31**(1): p. 88-102.
 40. Balvert, M., et al., *A framework for inverse planning of beam-on times for 3D small animal radiotherapy using interactive multi-objective optimisation.* Phys Med Biol, 2015. **60**(14): p. 5681-98.
 41. van Hoof, S.J., P.V. Granton, and F. Verhaegen, *Development and validation of a treatment planning system for small animal radiotherapy: SmART-Plan.* Radiother Oncol, 2013. **109**(3): p. 361-6.
 42. Mordant, P., et al., *Bioluminescent orthotopic mouse models of human localized non-small cell lung cancer: feasibility and identification of circulating tumour cells.* PLoS One, 2011. **6**(10): p. e26073.
 43. Shibamoto, Y., et al., *Influence of contrast materials on dose calculation in radiotherapy planning using computed tomography for tumors at various anatomical regions: a prospective study.* Radiother Oncol, 2007. **84**(1): p. 52-5.
 44. Liu, A.J., et al., *Effect of CT contrast on volumetric arc therapy planning (RapidArc and helical tomotherapy) for head and neck cancer.* Med Dosim, 2015. **40**(1): p. 32-6.
 45. Heydarheydari, S., N. Farshchian, and A. Haghparast, *Influence of the contrast agents on treatment planning dose calculations of prostate and rectal cancers.* Rep Pract Oncol Radiother, 2016. **21**(5): p. 441-6.
 46. Wichmann, J.L., et al., *Virtual monoenergetic dual-energy computed tomography: optimization of kiloelectron volt settings in head and neck cancer.* Invest Radiol, 2014. **49**(11): p. 735-41.
 47. Toepker, M., et al., *Can dual-energy CT improve the assessment of tumor margins in oral cancer?* Oral Oncol, 2014. **50**(3): p. 221-7.
 48. Grosse Hokamp, N., et al., *Improved visualization of hypodense liver lesions in virtual monoenergetic images from spectral detector CT: Proof of concept in a 3D-printed phantom and evaluation in 74 patients.* Eur J Radiol, 2018. **109**: p. 114-123.
 49. Scholtz, J.E., et al., *Non-linear image blending improves visualization of head and neck primary squamous cell carcinoma compared to linear blending in dual-energy CT.* Clin Radiol, 2015. **70**(2): p. 168-75.
 50. Albrecht, M.H., et al., *Assessment of an Advanced Monoenergetic Reconstruction Technique in Dual-Energy Computed Tomography of Head and Neck Cancer.* Eur Radiol, 2015. **25**(8): p. 2493-501.
 51. Grant, K.L., et al., *Assessment of an advanced image-based technique to calculate virtual monoenergetic computed tomographic images from a dual-energy examination to improve contrast-to-noise ratio in examinations using iodinated contrast media.* Invest Radiol, 2014. **49**(9): p. 586-92.
 52. McNamara, M.M., et al., *Multireader evaluation of lesion conspicuity in small pancreatic adenocarcinomas: complimentary value of iodine material density and low keV simulated monoenergetic images using multiphasic rapid kVp-switching dual energy CT.* Abdom Imaging, 2015. **40**(5): p. 1230-40.
 53. Yan, W.Q., et al., *Iodine Quantification Using Dual-Energy Computed Tomography for Differentiating Thymic Tumors.* J Comput Assist Tomogr, 2018. **42**(6): p. 873-880.

54. Ge, X., et al., *Comparative study of dual energy CT iodine imaging and standardized concentrations before and after chemoradiotherapy for esophageal cancer*. BMC Cancer, 2018. **18**(1): p. 1120.
55. Sun, K., et al., *Accuracy of Combined Computed Tomography Colonography and Dual Energy Iodine Map Imaging for Detecting Colorectal masses using High-pitch Dual-source CT*. Sci Rep, 2018. **8**(1): p. 3790.
56. Almeida, I.P., et al., *Monte Carlo proton dose calculations using a radiotherapy specific dual-energy CT scanner for tissue segmentation and range assessment*. Phys Med Biol, 2018. **63**(11): p. 115008.
57. Clark, D.P., et al., *In vivo characterization of tumor vasculature using iodine and gold nanoparticles and dual energy micro-CT*. Phys Med Biol, 2013. **58**(6): p. 1683-704.
58. Vaniqui, A., et al., *The impact of dual energy CT imaging on dose calculations for pre-clinical studies*. Radiat Oncol, 2017. **12**(1): p. 181.
59. Hoang Duc, A.K., et al., *Validation of clinical acceptability of an atlas-based segmentation algorithm for the delineation of organs at risk in head and neck cancer*. Med Phys, 2015. **42**(9): p. 5027-34.
60. Wardman, K., et al., *The feasibility of atlas-based automatic segmentation of MRI for H&N radiotherapy planning*. J Appl Clin Med Phys, 2016. **17**(4): p. 146-154.
61. Yang, J., et al., *Atlas ranking and selection for automatic segmentation of the esophagus from CT scans*. Phys Med Biol, 2017. **62**(23): p. 9140-9158.
62. Zaffino, P., et al., *Technical Note: plati-match mabs, an open source tool for automatic image segmentation*. Med Phys, 2016. **43**(9): p. 5155.
63. Zaffino, P., et al., *Multi atlas based segmentation: should we prefer the best atlas group over the group of best atlases?* Phys Med Biol, 2018. **63**(12): p. 12NT01.
64. Raudaschl, P.F., et al., *Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015*. Med Phys, 2017. **44**(5): p. 2020-2036.
65. Chapman, C.H., et al., *Deformable image registration-based contour propagation yields clinically acceptable plans for MRI-based cervical cancer brachytherapy planning*. Brachytherapy, 2018. **17**(2): p. 360-367.
66. Hardcastle, N., et al., *Accuracy of deformable image registration for contour propagation in adaptive lung radiotherapy*. Radiat Oncol, 2013. **8**: p. 243.
67. Li, X., et al., *Comprehensive evaluation of ten deformable image registration algorithms for contour propagation between CT and cone-beam CT images in adaptive head & neck radiotherapy*. PLoS One, 2017. **12**(4): p. e0175906.
68. Loi, G., et al., *Performance of commercially available deformable image registration platforms for contour propagation using patient-based computational phantoms: A multi-institutional study*. Med Phys, 2018. **45**(2): p. 748-757.
69. Thor, M., et al., *Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer*. Acta Oncol, 2011. **50**(6): p. 918-25.
70. Jin, R., et al., *Contour propagation for lung tumor delineation in 4D-CT using tensor-product surface of uniform and non-uniform closed cubic B-splines*. Phys Med Biol, 2017. **63**(1): p. 015017.
71. Liu, Y., et al., *Contour propagation using non-uniform cubic B-splines for lung tumor delineation in 4D-CT*. Int J Comput Assist Radiol Surg, 2016. **11**(12): p. 2139-2151.
72. Xing, L., E.A. Krupinski, and J. Cai, *Artificial intelligence will soon change the landscape of medical physics research and practice*. Med Phys, 2018. **45**(5): p. 1791-1793.
73. Han, X., *MR-based synthetic CT generation using a deep convolutional neural*

- network method. *Med Phys*, 2017. **44**(4): p. 1408-1419.
74. Wolterink, J.M., et al., *Generative Adversarial Networks for Noise Reduction in Low-Dose CT*. *IEEE Trans Med Imaging*, 2017. **36**(12): p. 2536-2545.
75. Zhang, K., et al., *Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising*. *IEEE Trans Image Process*, 2017. **26**(7): p. 3142-3155.
76. Yongshuai, G., et al., *Deconvolution-Based Backproject-Filter (BPF) Computed Tomography Image Reconstruction Method Using Deep Learning Technique*. arXiv, 2018.
77. Su, K.H., et al., *Machine learning-based dual-energy CT parametric mapping*. *Phys Med Biol*, 2018. **63**(12): p. 125001.
78. Hansen, D.C., et al., *ScatterNet: A convolutional neural network for cone-beam CT intensity correction*. *Med Phys*, 2018. **45**(11): p. 4916-4926.
79. Maier, J., et al., *Deep Scatter Estimation (DSE): Accurate Real-Time Scatter Estimation for X-ray CT Using a Deep Convolutional Neural Network*. *Journal of Nondestructive Evaluation*, 2018. **37**(3).
80. Nomura, Y., et al., *Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network*. *Med Phys*, 2019. **46**(7): p. 3142-3155.
81. Landry, G., et al., *Comparing Unet training with three different datasets to correct CBCT images for prostate radiotherapy dose calculations*. *Phys Med Biol*, 2019. **64**(3): p. 035011.
82. Roggen, T., et al., *Deep learning model for markerless tracking in spinal SBRT*, in ICCR. 2019.
83. Valdes, G., et al., *Clinical decision support of radiotherapy treatment planning: A data-driven machine learning strategy for patient-specific dosimetric decision making*. *Radiother Oncol*, 2017. **125**(3): p. 392-397.
84. van Wijk, Y., et al., *Decision Support Systems in Prostate Cancer Treatment: An Overview*. *Biomed Res Int*, 2019. **2019**: p. 4961768.
85. Walsh, S., et al., *Decision Support Systems in Oncology*. *JCO Clin Cancer Inform*, 2019. **3**: p. 1-9.
86. Widder, J., et al., *The Quest for Evidence for Proton Therapy: Model-Based Approach and Precision Medicine*. *Int J Radiat Oncol Biol Phys*, 2016. **95**(1): p. 30-6.
87. Landry, G., et al., *Improved dose calculation accuracy for low energy brachytherapy by optimizing dual energy CT imaging protocols for noise reduction using sinogram affirmed iterative reconstruction*. *Z Med Phys*, 2016. **26**(1): p. 75-87.
88. Bazalova, M., et al., *Dual-energy CT-based material extraction for tissue segmentation in Monte Carlo dose calculations*. *Phys Med Biol*, 2008. **53**(9): p. 2439-56.
89. Bazalova, M., et al., *Tissue segmentation in Monte Carlo treatment planning: a simulation study using dual-energy CT images*. *Radiother Oncol*, 2008. **86**(1): p. 93-8.
90. Aran, S., et al., *Applications of dual-energy CT in emergency radiology*. *AJR Am J Roentgenol*, 2014. **202**(4): p. W314-24.
91. Thieme, S.F., et al., *Dual Energy CT lung perfusion imaging--correlation with SPECT/CT*. *Eur J Radiol*, 2012. **81**(2): p. 360-5.
92. Vliegthart, R., et al., *Dual-energy CT of the heart*. *AJR Am J Roentgenol*, 2012. **199**(5 Suppl): p. S54-63.
93. Hidas, G., et al., *Determination of renal stone composition with dual-energy CT: in vivo analysis and comparison with X-ray diffraction*. *Radiology*, 2010. **257**(2): p. 394-401.
94. Almeida, I.P., et al., *Dual-energy CT quantitative imaging: a comparison study between twin-beam and dual-*

- source CT scanners. *Med Phys*, 2017. **44**(1): p. 171-179.
95. Gooding, M.J., et al., *Comparative evaluation of autocontouring in clinical practice: A practical method using the Turing test*. *Med Phys*, 2018. **45**(11): p. 5105-5115.
 96. Nikolov, S., et al., *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. *Arxiv*, 2018.
 97. Vaassen, F., et al., *Evaluation of measures for assessment of automatic organ-at-risk segmentation acceptability and time-savings*. *Physics and Imaging in Radiation Oncology*, 2019. **13**: p. 1-6.
 98. Alaei, P. and E. Spezi, *Imaging dose from cone beam computed tomography in radiation therapy*. *Phys Med*, 2015. **31**(7): p. 647-58.
 99. Ding, G.X., et al., *Image guidance doses delivered during radiotherapy: Quantification, management, and reduction: Report of the AAPM Therapy Physics Committee Task Group 180*. *Med Phys*, 2018. **45**(5): p. e84-e99.
 100. Dzierma, Y., et al., *Imaging dose and secondary cancer risk in image-guided radiotherapy of pediatric patients*. *Radiat Oncol*, 2018. **13**(1): p. 168.
 101. Oliver, J.A., et al., *Commissioning an in-room mobile CT for adaptive proton therapy with a compact proton system*. *J Appl Clin Med Phys*, 2018. **19**(3): p. 149-158.
 102. Eekers, D.B., et al., *The EPTN consensus-based atlas for CT- and MR-based contouring in neuro-oncology*. *Radiother Oncol*, 2018. **128**(1): p. 37-43.
 103. Kida, S., et al., *Cone Beam Computed Tomography Image Quality Improvement Using a Deep Convolutional Neural Network*. *Cureus*, 2018. **10**(4): p. e2548.
 104. Procz, S., et al., *X-ray and gamma imaging with Medipix and Timepix detectors in medical research*. *Radiation Measurements*, 2019. **127**.
 105. Ballabriga, R., et al., *Medipix3: A 64k pixel detector readout chip working in single photon counting mode with improved spectrometric performance*. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2011. **633**: p. S15-S18.
 106. Trojanova, E., et al., *Tissue sensitive imaging and tomography without contrast agents for small animals with Timepix based detectors*. *Journal of Instrumentation*, 2017. **12**(01): p. C01056-C0105

Valorization addendum

Background

The prevalence of cancer is a growing healthcare problem and a leading cause of death worldwide. The number of cancer deaths can be reduced through new developments in cancer diagnosis and cancer treatment. Cancer diagnosis techniques including laboratory tests, biopsies, and non-invasive imaging are all essential to detect nonmetastatic cancers at an early stage. Cancer treatment encompasses several techniques such as surgery, systemic therapy, and radiotherapy. Innovations in cancer diagnosis and treatment are required to increase cancer treatment efficacy. These innovations are also needed to reduce treatment induced side-effects during the treatment period. Radiotherapy induced side-effects are uncomfortable and potentially exist in the further lifespan of the patient.

Products, Innovations, and societal or commercial relevance

DirectDensity™ image reconstruction algorithm

Several techniques to improve cancer treatment and its workflow have been studied in this thesis. The first part of this thesis had a more clinically oriented focus with the commissioning of the DirectDensity™ image reconstruction algorithm in MV photon therapy. This study was part of a market evaluation phase study of the Siemens Healthineers Confidence RT pro dual-spiral CT scanner. The publication of this work gained interest from medical physicists since it describes the complete commissioning workflow of this novel CT image reconstruction algorithm for several treatment sites. After this study was published and presented at the largest European conference for radiotherapy and oncology, Siemens Healthineers estimated in July 2019 that over 100 scanners were licensed worldwide with the DirectDensity™ image reconstruction algorithm (private correspondence with Siemens Healthineers). Market evaluation phase studies are essential for industrial partners that do not have direct access to clinical data and clinical expertise. Industrial partners look at clinical collaborations as an added value, and for them, it is absolutely required to evaluate the usability of the algorithm or device on more clinically oriented criteria. **Chapter 3** of this thesis contributed to the market evaluation phase of Siemens Healthineers's DirectDensity™ image reconstruction algorithm.

VOXSI simulation software

The research carried out at Maastricht and at other institutes worldwide has resulted in a large number of publications which have shown the benefits of dual-energy CT in radiotherapy. The acquired knowledge from past years and the best performing algorithms in dual-energy CT available in literature were combined in a free and open-

access simulation software package for single- and dual-energy CT (VOXSI). VOXSI has been made publicly available for online download⁶. After development, this simulation software is being actively used for teaching and research purposes by several universities and companies. Later, open-source access and development support was provided to two international research institutions which are currently working on extension packages for proton radiography and advanced noise modelling. Other free open-access simulation software is often complicated to install and does not provide built-in image reconstruction and analysis software. Our simulation software directly enables advanced single-energy CT and dual-energy CT post-processing which is currently not available in any other software simulation package. VOXSI has been programmed in an object-oriented and a modular way, which enables the easy implementation of future science and technology inside the software.

Monte Carlo simulations

Monte Carlo simulation software is often used in radiotherapy to comprehend systems, experiments, or devices that deal with ionizing radiation. In this thesis, a Monte Carlo simulation model was developed from a novel imaging ring system on rails in collaboration with an industrial partner (medPhoton GmbH, Salzburg) (**Chapter 6** and **Chapter 7**). This imaging ring simulation model was actively used to explore image quality improvements and was also used to investigate clever artificial intelligence methods that are able to correct for photon scatter in head and neck cancer patients. Although further investigations are needed to investigate the scatter removal performance (**Chapter 7**) in larger body sites and for other imaging spectra, this algorithm showed superior results compared to other advanced corrections in smaller body sites and has potential to be implemented in clinical software. All research in this project was part of an Eurostars framework between Maastricht University and medPhoton GmbH.

Automatic image segmentation software

The second part of this thesis was more focused on full-automatic image segmentation in preclinical research. Similar methods were also applied in **Chapter 5** to clinical dual-energy CT data in collaboration with OncoRay in Dresden. To make significant progress in curing many diseases it is necessary to perform meaningful preclinical research. Biological oriented researchers have limited their preclinical image analysis in the past to simplified analysis methodologies because extensive imaging analysis was extremely time consuming. I felt there was a great need to develop powerful technology that allows performing complex animal studies that require preclinical imaging. Two different methods were developed in **Chapter 8** and

6 <https://voxsi.weebly.com>

Chapter 9 to segment volumes-of-interest on micro cone-beam CT scans of mice. This work highlighted the great interest of preclinical researchers in automatic segmentation algorithms, but also the need of it. Dedicated automatic image segmentation algorithms improve the workflow speed and facilitate the preclinical workflow significantly. Furthermore, they have the potential to reduce the number of lab animals used in longitudinal preclinical research because the automatic analysis tools could replace existing methods which required sacrificing a cohort of animals.

Not every preclinical institution has access to automatic image segmentation algorithms or has the knowledge to build software themselves. Based on the research described in **Chapter 8** and **Chapter 9** of this thesis, a grant proposal was submitted and funded (OMEGA) to develop commercial automatic contouring software with deep learning to reduce animal experiments, as a collaboration of Maastricht University and a commercial spinoff company. In the OMEGA project, tools and software will be developed to broaden the use of preclinical CT segmentation software. These tools will be available and advertised to the preclinical research community as a commercial product and service.



Acknowledgements

I would first like to acknowledge my supervisor Prof. Dr. Frank Verhaegen together with my co-promoters. **Frank**, thank for your guidance, your enthusiasm, your novel research ideas, and your always open door through every stage of my PhD. I admire your ability to inspire people, to listen to issues, and adapt to change. I really couldn't wish for a better supervisor during my PhD and I hope our collaboration can continue in the future.

Mark, *señor*, we were only Maastricht-neighbors for a short time period in the beginning of my PhD trajectory, but in this short period, you taught me so many things and inspired me in so many ways. I couldn't even find words for it. Thank you for all your patience, help, and good sense of Irish humor. Like you would have said: 'cheers man!'.

Gabriel, thank you for your help in the last years of my PhD. I admired how you dealt with stressful situations, and how you took time to help everyone, even when there were many more students in the famous and almost-patented Gabriel's question queue.

Stefan, although our career paths split after my master's thesis, you actually contributed significantly to this work without being aware, I think. Even today, I still use many skills you taught me as a master's student. I forgot to apologize that I was bothering you almost every 30 minutes with a new question in the beginning of my MSc internship, but this was only because you solved my last question so quickly. If Matlab was an official language, you were a native speaker! Jose, thank you for being my Spanish teacher and my Maastricht party partner in crime. I enjoyed every single night out, and of what remains, I only have good memories to a wonderful time. Good luck with your new position in Australia! Hasta pronto amigo! Ana, I admired how you finished your PhD in record speed, working long days, but mostly evenings and nights, and I am pretty sure you will have the career you deserve! Wouter v. Elmpt, sorry that I couldn't find a better analogy, but I would like to compare your office with an escape room. Not that I wanted to escape so quickly, no ..., entering your office was always like going on a journey. I knew when I entered, but I never knew when I would come out, and when I did, I was full of new insights and new ideas. Thank you!

Further, I want to acknowledge all colleagues I actively worked with in the past years: Behzad, Cecile, Céline, Colien, Daniëlle, Femke, Frank D., Gloria, Ilaria, Isabel, Jan, Karin, Michel, Murillo, Nadine, Ramon, Relinde, Rosa, Richard, Shane, Sonia, Teun, Timo, and Wouter vd. Worp. Over the past years, it was also fun to meet many people from other research groups: Aniek, Evelyn, Janita, Jurgen, Karen, Ralph, Ruben, Sean, Sebastian, Turkey, and Yvonka.

Most significantly I am grateful to my parents Alain and Veerle, my grandparents, my sister Jorien, and my brother in law David. They always encouraged and supported me in pursuing my interests. Bedankt!



Curriculum Vitae

Brent van der Heyden was born on October 19th, 1994 in Sint-Niklaas. He completed his general secondary education in 2012 at GTI Beveren and started studying applied nuclear engineering technology in Diepenbeek. In June 2016, Brent attained his master's degree (cum laude) in medical-nuclear engineering technology at the joint scholar program of Hasselt University and the Catholic University of Leuven. During his MSc internship at Maastrro in Maastricht, Brent modelled the adverse effects of respiratory motion in mice on preclinical precision irradiations. In July 2016, Brent started his PhD at the physics research group in Maastrro under the supervision of Prof. Dr. Frank Verhaegen. During his PhD, he mainly worked on computed tomography applications for (pre)clinical radiotherapy and was also involved in several other studies related to artificial intelligence, bioluminescence imaging, and dose-guided radiotherapy.





Publication list

List of peer-reviewed publications

Published

B. van der Heyden, G. Fonsceca, M. Podesta, I. Messner, H. Deutchmann, P. Steinger, and F. Verhaegen. “Modelling of the focal spot intensity distribution and the off-focal spot radiation in kilovoltage X-ray tubes for imaging” *Physics in Medicine and Biology* 2019

B. van der Heyden*, W. van de Worp*, A. van Helvoort, J. Theys, A. Schols, R. Langen*, and F. Verhaegen*. “Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network” *Journal of Applied Physiology* 2019

B. van der Heyden*, I. Almeida*, G. Vilches-Freixas, C. van Beveren, A. Vaniqui, C. Ares, K Terhaag, G. Fonseca, D Eekers, and F Verhaegen. “A comparison study between single- and dual-energy CT density extraction methods for neurological proton Monte Carlo treatment planning” *Acta Oncologica* 2019

B. van der Heyden*, P. Wohlfahrt*, E. Troost, K. Terhaag, D. Eekers, C. Richter, and F. Verhaegen. “Dual-energy CT for automatic organ-at-risk segmentation in brain-tumor patients using a multi-atlas and deep-learning approach” *Scientific Reports* 2019

F. Vaassen, C. Hazelaar, A. Vaniqui, M. Gooding, **B. van der Heyden**, R. Canters, W. van Elmpt. “Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy” *Physics and Imaging in Radiation Oncology* 2019

B. van der Heyden, M. Podesta, D. Eekers, A. Vaniqui, I. Almeida, L. Schyns, S. van Hoof, and F. Verhaegen. “Automatic multi-atlas based organ at risk segmentation in mice” *British Journal of Radiology* 2018

B. van der Heyden, L. Schyns, M. Podesta, A. Vaniqui, I. Almeida, G. Landry, and F. Verhaegen. “VOXSI: a voxelized single- and dual-energy CT scenario generator for quantitative imaging” *Physics and Imaging in Radiation Oncology* 2018

J. Yang, H. Veeraghavan, SG. Armato, K. Farahani, JS. Kirby, J. Kalpathy-Kramer, W. van Elmpt, A. Dekker, X. Han, X. Feng, P. Aljabar, B Oliveira, **B. van der Heyden**, L. Zamborg, D. Lam, M. Gooding, and GC. Sharp. “Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017” *Medical Physics* 2018 (Editor’s choice)

L. Schyns, D. Eekers, **B. van der Heyden**, I Almeida, A Vaniqui, and F. Verhaegen. “Murine vs human tissue compositions: implications of using human tissue compositions for photon energy absorption in mice” *British Journal of Radiology* 2018

A. Vaniqui, **B. van der Heyden**, I. Almeida, L. Schyns, S. van Hoof, and F. Verhaegen. “On the determination of planning target margins due to motion for mice

lung tumours using a four-dimensional MOBY phantom” British Journal of Radiology 2018

A. Vaniqui, L. Schyns, I. Almeida, **B. van der Heyden**, M. Podesta, and F. Verhaegen. “The effect of different image reconstruction techniques on pre-clinical quantitative imaging and dual-energy CT” British Journal of Radiology 2018

I. Almeida, A. Vaniqui, L. Schyns, **B. van der Heyden**, J. Cooley, T. Zwart, A Langenegger, and F. Verhaegen. “Exploring the feasibility of a clinical proton beam with an adaptive aperture for pre-clinical research” British Journal of Radiology 2018

I. Almeida, L. Schyns, A. Vaniqui, **B. van der Heyden**, G. Dedes, AF. Resch, F. Kamp, JD. Zindler, K. Parodi, G. Landry, and F. Verhaegen. “Monte Carlo dose calculations using a radiotherapy specific dual-energy CT scanner for tissue segmentation and range assessment” Physics in Medicine and Biology 2018

A. Vaniqui, L. Schyns, I. Almeida, **B. van der Heyden**, S. van Hoof, and F. Verhaegen. “The impact of dual energy CT imaging on dose calculations for pre-clinical studies” Radiation Oncology 2017

B. van der Heyden, M. Öllers, A. Ritter, F. Verhaegen, and W. van Elmpt. “Clinical evaluation of a novel CT image reconstruction algorithm for direct dose calculations” Physics and Imaging in Radiation Oncology 2017

B. van der Heyden, S. van Hoof, L. Schyns, and F. Verhaegen. “The influence of respiratory motion on dose delivery in a mouse lung tumour irradiation using the 4D MOBY phantom” British Journal of Radiology 2016

Submitted

B. van der Heyden^{*}, M. Uray^{*}, G. Fonseca, P. Huber, D. Us, I. Messner, A. Law, A. Parri, N. Reisz, I. Rinaldi, G. Vilches-Freixas, H. Deutschmann, F. Verhaegen^{*}, and P. Steininger^{*}. “A Monte Carlo based scatter removal method for non-isocentric cone-beam CT acquisitions using a deep convolutional autoencoder” Submitted work.

In preparation

B. van der Heyden et al. X-ray tube and jaw modulation techniques for imaging dose reduction in non-isocentric cone-beam CT trajectories.

* equal contribution

