2D patient setup verification in MRI-only radiotherapy versus current CT-based verification

Manija Ghafory

Herlev Hospital

Kongens Lyngby 2013-23
IMM-MSc-2013-23

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk IMM-MSc-2013-23

Summary (English)

Radiation therapy (RT) is one of the most common treatment method for cancer patients. The purpose of RT is to cure the patient through ionizing radiation. This requires a treatment planning process with high accuracy. The current treatment planning is based on a computed tomography (CT) scan which contains information about the electron densities, which are required for dose calculation. The CT scan is also important for 2D patient setup verification. RT based on magnetic resonance imaging (MRI) has proved advantages compared to the CT due to e.g. better delineation of tumour volume and organ at risks. The aim of this study is to investigate the possibility of using MRI for 2D patient setup verification.

Data from four palliative patients receiving cranial RT was used in this study. The patients were scanned with 1 Tesla open MRI-system and MRI ultrashort echo-time (UTE) sequence scans were required. The Markov random field (MRF) segmentation method was used to classify each MRI UTE sequence data into air, soft tissue and bone and created a substituted CT (sCT) scans. sCT bone digital reconstructed radiographs (DRRs) were generated from the sCT scans and CT bone DRRs were generated from the planning CT scans. Manual match of OBIs on both CT DRRs and sCT DRRs were performed in Offline Review Eclipse V.10 (Varian Medical System). A 2D lateral and frontal match was performed by five radiation therapists (RTTs). A statistical evaluation was made of whether there is a significant difference in 2D patient setup verification when performing matching on sCT generated DRRs as compared to CT generated DRRs.

The MRF segmentation facilitated creating sCT scan and generated bone DRRs.

A significant difference between the sCT and CT generated DRRs was seen in longitudinal (lateral) direction, vertical direction and pitch (rotation) when performing 2D lateral match. For the frontal match a significant difference was observed in rnt (rotation) whereas longitudinal (front) and lateral directions were non-significant.

This study showed that treatment planning solely based on MRI is a feasible alternative to current CT based on treatment planning due to 2D setup verification.

Summary (Danish)

Strålebehandling er en af de mest almindelige behandlingsmetoder til kræftpatienter. Formålet med strålebehandling er at helbrede patienten ved hjælp af ioniserende stråling. Dette kræver en nøjagtig planlægningsproces af behandlingen. Den nuværende planlægning er baseret på en Computed tomography (CT) skanning, som indeholder oplysninger om elektron densiteter, som er nødvendig for doseberegning. CT skanningen bruges også til 2D patient setup verifikation. Strålebehandling baseret på magnetisk resonans skanninger (MR) har vist fordele i forhold til CT bl.a. i forhold til optegning af tumor volume og risiko organer. Formålet med dette studie er at undersøge muligheden for at anvende MR til 2D patient setup verifikation.

Data fra fire palliative patienter som har fået kraniel strålebehandling blev anvendt. Patienterne blev skannet med 1 Tesla open MRI - system og MR ultrashort echo-time (UTE) sekvens skanninger var optaget. Markov random field (MRF) segmenterings metode blev anvendt til at klassificere hvert MRI UTE sekvens data til hhv. luft, blødt væv og knogle og skabte sCT skanninger. sCT knogle digital reconstructed radiographs (DRRs) blev genereret fra sCT skanninger og CT knogle DRR blev genereret fra planlægnings CT-skanninger. Manuel match af 2D OBI på både CT DRR og sCT DRR blev udført i Offline Review Eclipse V.10 (Varian Medical System). En 2D lateral og frontal match blev udført af fem radioterapeuter. En statistisk evaluering af, om der er en signifikant forskel i 2D patient setup verifikation ved udførelse af sCT genereret DRR match med CT genereret DRR blev foretaget.

MRF segmenteringen muliggjorde genereringen af sCT skanninger og knogle DRR. En signifikant forskel mellem sCT og CT genereret DRR blev set i den

laterale retning, vertikale retning og pitch (rotation) ved udførelse af 2D lateral match. For det frontale match en signifikant forskel blev observeret i RNT (rotation), hvorimod de longitudinale (frontale) og laterale retninger var ikke-signifikante.

Studiet viste, at planlægningen af behandling udelukkende baseret på MRI er et realistisk alternativ til den nuværende CT baseret behandlings planlægning for 2D setup patient verifikation.

Preface

This master project has been carried out in close co-operation with the Department of Oncology at Copenhagen University Hospital, Herlev Hospital and the Department of Informatics and Mathematical Modelling at the Technical University of Denmark. This is the final project for receiving Master of Science degree in Engineering (Medicine and Technology) at Technical University of Denmark and Copenhagen University, the Faculty of Health Science. This MSc thesis is carried out in the time period September 3rd, 2012 - March 18th, 2013 and credited 35 ECTS point.

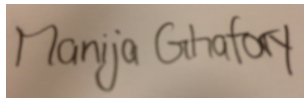
Supervisors:

Jens Edmund, PhD, DABR
Department of Oncology
Copenhagen University Hospital
Herlev Hospital

Rasmus Larsen, Professor
Department of Informatics and Mathematical Modelling
Technical University of Denmark

Knut Conradsen, Professor
Department of Informatics and Mathematical Modelling
Technical University of Denmark

Lyngby, 18-March-2013-23



Manija Ghafory

Acknowledgements

I would like to thank my supervisors PhD, DABR Jens Edmund from Department of Oncology at Copenhagen University Hospital, Herlev Hospital, Professor Rasmus Larsen and Professor Knut Conradsen, both from the Department of Informatics and Mathematical Modelling, Technical University of Denmark for their helpful guidance throughout the project.

I want to give a special thanks to PhD student, Hans Martin Kjer for helping me with data processing and guidance in software programming.

Thanks to PhD student Mark Lyksborg for guidance in software programming.

Also thanks to the five radio therapists at the Department of Oncology at Copenhagen University Hospital, Herlev Hospital, for using their expertise and knowledge in data analysis.

Contents

List of Figures

List of Tables

List of Acronyms

ANOVA	Analysis of variance
BEV	Beam's-eye-view
CBCT	Cone beam computed tomography
CT	Computed tomography
DF	Degrees of freedom
DICOM	Digital imaging communications in medicine
DRR	Digital reconstructed radiograph
dUTE	dual Ultrashort echo-time
FID	Free induction decay
G-G	Greenhouse-Geisser
H-F	Huynh-Feldt
H-F	Huynh-Feldt
LINAC	Linear accelerator
LSD	Least significant difference
MANOVA	Multivariate analysis of variance
MRF	Markov random field
RT	Radiation therapy
MRI	Magnetic resonance imaging
OBI	On-Board imager
HU	Hounsfield unite
PET	Positron emission tomography
PIPSpro	Portal image processing system
PTV	Planning target volume
REV	Room's-eye-view
RF	Radio frequency
RT	Radiation therapy
RTT	Radio therapist
sCT	Substituted CT
T-box	Training box
TE	Echo time
TPS	Treatment planning-system
UTE	Ultrashort echo-time

Chapter 1

Introduction

Radiation therapy (RT) is a treatment method for many cancer types and it is one of the most common cancer treatments. The cancer cells can permanently be destroyed, if the dose is high enough, but adjacent healthy cells can also be damaged. This can lead to several side effects. The aim of RT is to cure the patient through ionizing radiation. This requires a treatment planning process and treatment delivery with high accuracy.

Modern RT requires 3D based images that contain information about the patient's anatomy to delineate tumour volume and organ at risks (OARs) delineation and verify patient setup prior to treatment delivery. The images can be obtained with modern imaging techniques, such as computed comography (CT) and magnetic resonance imaging (MRI) [20].

The current practice of treatment planning is based on a CT scan of the patient which contains information about the electron densities. Images are required when calculating the dose plans. The tumour volume and OARs are delineated on CT images. Often RT requires multimodality imaging with MRI and CT for a more accurate tumour volume and OARs delineation. This kind of treatment process requires a long workflow, as shown in figure 1.1. The CT scan is also used to generate digital reconstructed radiographs (DRRs) of bone structures which are used for patient setup verification.

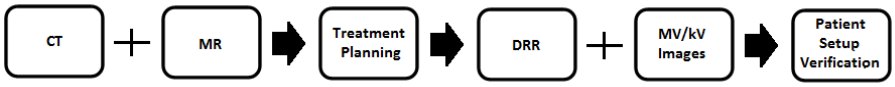


Figure 1.1: Workflow of the current process of radiation treatment planning using multi-modality imaging techniques.

Several studies show how to facilitate the use of an MR imaging technique for the process of radiation therapy planning [5, 9]. The MR imaging technique has various advantages e.g. better delineation of the tumour volume and OARs, since these structures can be adequately identified in MR images [2]. In addition, treatment planning solely based on MRI requires less workflow compared to RT using multi-modality techniques.

The main focus of this study is to replace the current CT-based RT due to patient setup verification with MRI by creating a substituted CT, so-called sCT. Previous studies have shown that a sCT can be created from an MR scan by the use of dual ultrashort echo time (dUTE) sequence. The output from MR dUTE sequence can be segmented into air, soft tissue, and compact bone by performing different segmentation strategies [10, 9].

In this study, four palliative patients receiving cranial RT were scanned with the MR dUTE sequence scanner. A Markov random field classification strategy was performed to generate the sCTs. The possibility of using a sCT scan in the process of radiation therapy planning was investigated by performing a clinical evaluation in 2D patient setup verification.

1.0.1 Objective

The main aim of this study is to carry out a clinical evaluation using 2D Digital Reconstructed Radiograph (DRR) patient images from two image modalities for patient setup verification. The modalities are a CT scan and a sCT scan, respectively. It is investigated whether sCT DRRs can substitute the current CT DRRs used for 2D patient setup verification.

To accomplish the main aim of this study, the objectives below were pursued.

- Perform a Markov random filed classification to generate a sCT scan.
- Produce a method i.e. a software program to 2D make manual matching of OBIs on both DRRs generated from a CT and a sCT scan.
- Recruit experienced radio therapists (RTTs) to match kV images with DRRs generated from both CT scan and sCT scan. This will be performed for each patient fraction in a random order and the sCT and CT generated DRRs will be blinded.
- Make a statistical evaluation of whether there is a significant difference in 2D setup verification of patients when performing the matching on sCT generated DRRs as compared to normal CT based DRRs.

1.0.2 Previous Work

Two related works are preferred in this study, Kjer Master Thesis [10] and Buhl et al. [2], respectively.

Kjer [10] investigated the possibility of visualising compact bone in MRI images when using ultra short echo times (UTE) sequence. The data in this study was based on a calf knee and single patient head anatomy.

Different tissue classification strategies were performed to create a substituted CT scan, so-called sCT scan. This was done by segmenting the tissues appearing in both calf knee MRI UTE image and patient head anatomy MRI UTE image.

The results from the different classification strategies showed that Markov random field method has obtained an overall best result both on knee UTE MRI data and patient UTE MRI data. Therefore the use of MRF classification method is preferred in this study.

Buhl et al. [2] investigated 3D/3D MRI-CBCT automatching on brain tumours for online setup verification. Two experiments were made in this study, a multimodality phantom and clinical experiment. The aim of the phantom experiment was investigated whether it is feasible to perform online 3D/3D MRI-CBCT automatch and compared to the 3D/3D CT-CBCT automatching. The clinical experiment were included three patients receiving RT for malignant brain tumours. 18 CBCT were matched both with CT and MRI as a reference.

The result based on t-test from the phantom experiment showed no significant difference between MRI-CBCT and CT-CBCT for the vertical and the lateral directions, but a significant difference was seen for the longitudinal direction, and MRI-CBCT obtained the best automatch. Therefore it was concluded that it is feasible to perform 3D/3D MRI-CBCT automatching for online patient verification. The results from clinical experiment showed no difference > 3 mm for longitudinal and lateral direction. For the vertical direction up to 2 mm were observed. The mean and standard deviations showed that MR-CBCT performed not significantly worse than CT-CBCT. Buhl et al. study showed that it is possible to conduct 3D/3D MRI-CBCT automatching for patient online setup verification for brain RT.

The work of Buhl et. al [2] is used as an inspiration for 2D patient setup verification in this study.

Chapter 2

Theory

2.1 Radiation Therapy Planning Process

Cancer treatment with radiation therapy requires a planning process with high accuracy. The following processes need to be carried out before delivery of the treatment: establish the patient's treatment position, construct the patient repositioning and immobilization, image acquisition, delineate tumour volume and organ at risks, beam design, calculate dose for the treatment, evaluate dose plan, verify patient position and plan [20].

2.1.1 Immobilization

Treatment with radiation therapy (RT) requires accurate, stable, reproducible patient positioning throughout the treatment at each fraction before the treatment is delivered to the patient. The patient must be informed about the importance of remaining still during the treatment in order to obtain the planned beam direction to irradiate the planning target volume (PTV) [16, 6].

Immobilization devices are used to minimize patient movements and to reproduce patient positioning shown in figure 2.1. The immobilization devices are

applied to obtain a fixed position throughout the treatment, e.g. bolus bags for general support vacfix, thermochemical polystyrene headrests and knee supports, vacuum bags for breast and pelvic treatment etc. [23].

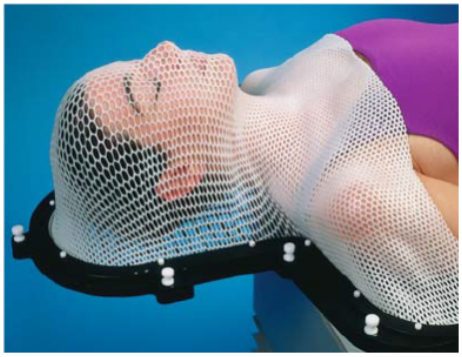


Figure 2.1: Immobilization device for head and neck. This image is modified from [20].

Most immobilization devices are made of carbon which enables the beam to pass through the material without disturbing the dose distribution. In addition, markers are placed on the patient's skin and on the immobilization device to serve as fiducial marks for the treatment setup verification [23].

2.1.2 Image Acquisition

Different imaging modalities are applied in radiation therapy, such as computed tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI). Often multi-modality is required to identify the location and size of the gross tumour volume and OARs. CT imaging technique is the golden standard for the radiation therapy planning process.

2.1.3 Treatment Planning

The RT workflow description in this section is mainly based on Prince et al. [20]. The planning CT data set is used to delineate the tumour volume and OARs and in some cancer treatments PET and MRI are also used for delineation. This procedure is performed by a radiation oncologist and a radiologist. The delineations of target volume and OARs are drawn manually slice by slice and

in some cases OARs with distinct boundaries are contoured automatically, e.g. lungs, body surface and bone.

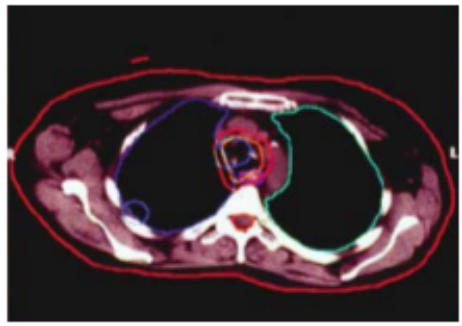


Figure 2.2: CT axial slice of thorax where the tumor, target volume and organ at risks are delineated. The image is modified from [20].

Figure 2.2 illustrates a CT scan of lungs where contours are drawn around the target volume and adjacent tissues on a slice by slice basis. Treatment Planning-System (TPS) is an effective tool for the treatment planners and oncologists to delineate these structures. Designing beam arrangement and field apertures is the next step in the treatment planning process. The treatment delivery is established on the basis of patient clinical protocol, diagnostic group and the location of the tumour. A 3D TPS beam's-eye-view (BEV) is an important display tool to identify the best collimator, gentry and couch angle to radiate the target volume and prevent radiation of OARs. A 3D TPS room's-eye-view (REV) display is also a powerful tool enabling planner to simulate any arbitrary viewing location within the treatment room. REV display helps the treatment planner to increase accuracy of overall beam arrangement geometry and positioning of the isocenter.

The positional accuracy of the patient is a very important treatment factor, and it must be obtained before treatment delivery to the tumour. Digital Reconstructed Radiographs (DRRs) are generated from CT scans. DRR of bones is used for verification of the patient positioning at the linear accelerator (LINAC).

The dose calculation is based on an algorithm that accurately calculates the dose for a patient. The optimization for dose distribution is achieved in TPS, which is based on maximization of dose to target volume and minimization of dose to OARs.

Documentation for plan implementation is complete when the treatment plan is designed, evaluated and approved. A parameter and plan check is performed by the physicist.

2.1.4 Position Verification and Treatment Delivery

The delivery of the treatment to the patient is acquired at LINAC. Prior to the treatment delivery at each fraction, it is important to reproduce the patient positioning.

The on-board imager (OBI) in the LINAC is used to acquire orthogonal kV images which are applied as a verification tool for patient positioning. The acquired kV images are produced by using high resolution x-rays. The kV images or OBIs improve the dose delivery to the target volume by minimizing patient movement during treatment [1].

The two orthogonal, frontal and lateral OBIs are matched with the corresponding planning CT generated DRRs at the same angle. High dense material can be seen in the images and the manual registration/matching of DRRs and OBIs are therefore based on the bone structures. The match of OBIs with DRRs allows small couch adjustments to verify the planned positioning of the patient. When planned patient positioning is verified, the treatment of the patient can commence.

2.2 Imaging Modalities

2.2.1 Computed Tomography

Computed tomography (CT) is based on x-ray, which generates 2D projection images of the body. The x-ray projections are obtained as the x-ray tube rotates around the patient and different beam angles pass through the patient. The beams penetrate into the body and the attenuation of the beam depends on tissue type. The detectors are placed at the opposite side of the x-ray tube and measure the intensity of the attenuated beams. The detectors will convert the intensities into signals as CT raw data. The CT scanner reconstructs the value of attenuation μ at each pixel of raw CT data within a cross section [19, 8].

The filtered back-projection is the most used algorithm to create a reconstructed CT image which is a grey tone image. The CT numbers are computed from linear attenuation coefficients at each pixel. CT values, also known as Hounsfield values can be defined as follows [19, 8]:

$$hu = \frac{\mu_m - \mu_w}{\mu_w} \times 1000 \tag{2.1}$$

where μ_m refers to linear attenuation coefficient within voxels and μ_w represents the linear attenuation coefficient for water at the same photon energy average of spectra.

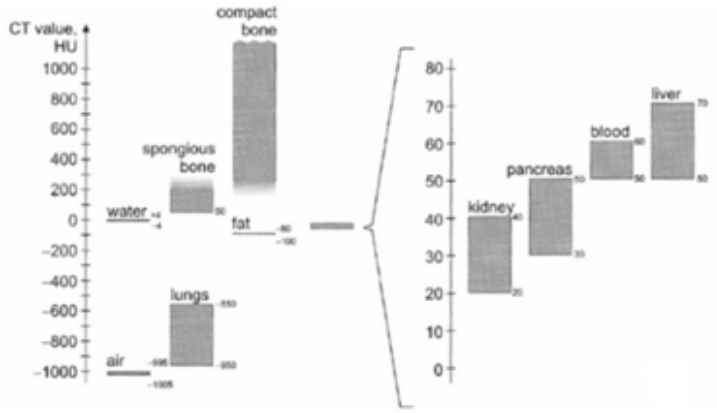


Figure 2.3: Hounsfield values of different tissues. This image is modified from [8].

Figure 2.3 illustrates Hounsfield values of different tissues, where water has 0

HU by definition and air has -1000 HU. The largest Hounsfield values found in the body are bones, where $hu = 1000$ HU for average bone.

2.2.2 Digital Reconstructed Radiograph

In RT, the planning CT data is used for the verification of patient positioning prior to treatment delivery. 2D x-ray images of a patient from a given angle are calculated using the planning CT data. These 2D x-ray images are called digital reconstructed radiographs (DRRs) [15].

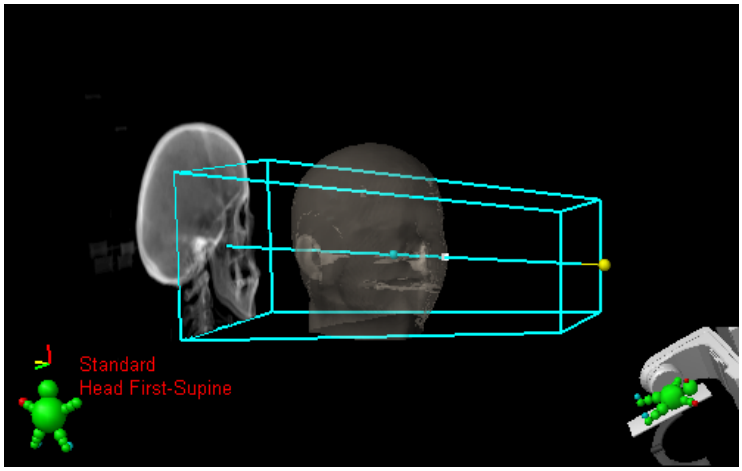


Figure 2.4: Generation of DRR.

DRRs are generated from the planning CT data using TPS as shown in figure 2.4. The generated CT DRRs are acquired for the patient setup verification at the LINAC.

2.2.3 Magnetic Resonance Imaging

MRI is a powerful non-invasive imaging modality. MR is frequently used for diagnostic investigations of for example the central nervous system and musculoskeletal system.

MRI is a useful imaging technique in radiation therapy since it provides an important added feature to CT imaging during delineation of target volume and organs at risk.

Tissues with short T2 relaxation, such as cortical bone, produce a very low or no signal since the signal decays rapidly after excitation when using conventional MRI sequences which typically use an echo time (TEs) of several milliseconds or longer. Therefore, the tissues with short T2 cannot be visualized since the tissues appear dark and this makes it difficult to separate bone from air [22]. The use of ultra short echo-time (UTE) pulse sequence allows detection of signals

Tissues	Mean T2
Ligaments	4-10 ms
Periosteum	5-11 ms
Cortical bone	0.42-0.5 ms

Table 2.1: The table shows the mean T2 of some tissues at 1.5 Tesla [22].

from tissues with short T2, such as cortical bone (shown in table 2.1). The echo time is defined as the time from first excitation to signal readout. To reduce T2 signal loss, the duration of the RF excitation pulse and the echo time must be minimized. The use of a small flip angle, allows the radio frequency (RF) pulse to be kept below 100 μs. There is no time to recall and sample an echo, therefore the UTE sequence samples the free induction decay (FID) rather than a gradient echo. The contrast in the sampled FID signal images is controlled by T2*, therefore susceptibility artefacts can appear between air and soft tissue interfaces [22, 9, 10, 21].

2.2.3.1 UTE sequence imaging

As mentioned in the previous section, it is difficult to separate bone and air with the conventional MRI sequence. By applying UTE sequence the signals from tissues with short decay can be detected. There are several imaging strategies with UTE. One of the UTE imaging strategies uses two different or dual echo times (dUTEs) to measure signals from cortical bone and other soft tissues. A second echo is obtained shortly after the first [10, 9].

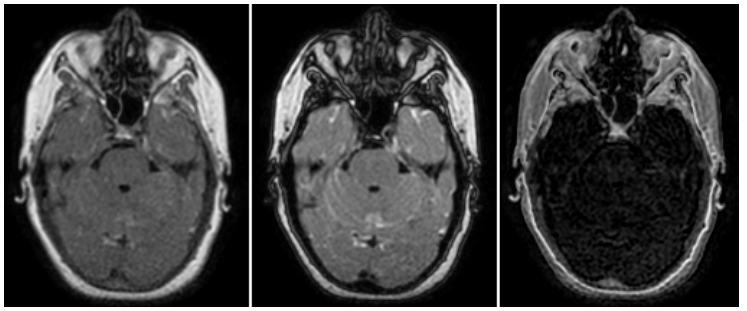


Figure 2.5: The figure shows axial images of the patient brain included in this study acquired with UTE sequence. The image to the left refers to the first echo (Echo 1), the image in the middle shows the second echo (Echo 2) and the image to the right is a subtraction image.

Figure 2.5 illustrates the first echo, second echo and subtraction images in axial plane. Tissues with short T2 lose a lot of signal in the time between the echoes, and therefore they appear bright. Cortical bone is present in the Echo 1 image with a high signal, although still very weak compared to the signal from soft tissue, but there is no signal from cortical bone in Echo 2. In the subtraction image, cortical bone appears with a high intensity as a thin bright contour of the bone. Soft tissues are present in both Echo 1 and Echo 2 with very similar intensity and are therefore almost absent in the subtraction image. Air has a very low intensity and cannot be seen in either Echo 1 or Echo 2.

2.2.4 Markov Random Field Segmentation

The MR data which is obtained from an MRI scanner with dUTE sequence consists of two image volumes, first echo and second echo, respectively. A segmented scan can be created where all voxels are assigned to a group that represents specific HU. The three main groups are considered as air (-1000 HU), bone (500-2000 HU) and soft tissue (0 HU), respectively [10].

Markov Random Field (MRF) is used to model a prior probability of context dependent patterns such as image voxel. This is obtained from mutual influences among entities using conditional MRF distributions. MRF prefers its own class of patterns by associating them with larger probabilities than other pattern classes [12].

In this study, MRF is used as a voxel classifier where each voxel is assigned into one of k different classes. This is referred to as label volume where class or

label is corresponded to a type of tissue that appears in the UTE sequence MR image [10]. MRF changes the posterior probability of each pixel by looking at the surrounding neighbourhood.

The task is to assign each voxel i in a volume into one of k different tissue classes. Using Bayes's rule the posterior probability can be estimated for a given voxel with intensity $x = [x_1, x_2]$.

$$P(k|x) = \frac{P(x|k)P(k)}{P(x)} \tag{2.2}$$

where $P(x|k)$ is the conditional probability and it states the probability that intensity value x is in the $k'th$ class and $P(k)$ is the prior probability and $P(x)$ is a normalizing function.

The prior probability used where voxels with the same label are clustered spatially throughout the image.

$$P(k) = \frac{1}{z} exp(-E(k)) \tag{2.3}$$

where z is a normalizing constant and $E(k)$ is an energy function. If the energy function $E(k) \longrightarrow 0$ a high number of neighbouring voxels are spatially clustered in the same area and if energy function $E(k) \longrightarrow \infty$ a low number of neighbouring voxels are spatially clustered [10, 12].

Equation 2.4 facilitates the definition of the prior which takes the local neighbouring of voxels into account [11, 10];

$$P_i(k) = \frac{\pi_k exp(-\beta \sum_{j \in \Re 1}(1 - q_i(k)))}{\sum_{k'} \pi_{k'} exp(-\beta \sum_{j \in \Re 1}(1 - q_i(k')))} \tag{2.4}$$

where π_k is class prior constant from the Bayes classifier, β is a constant that predicts the influence of local neighbouring of voxels, $q_j(k)$is the current posterior probability of voxel j belonging to class k and the summation is the summing of probabilities of voxels that are not belonging to class $(1 - q_j(k))$ in a local neighbouring $\Re_1$, [10, 11].

The MRF classifier is performed to classify each MRI UTE sequence data set into air, soft tissue and compact bone. Six to seven classes are considered to classify the tissues that appear in MR image and an sCT scan is created.

2.3 Image Registration

Image registration is a process which aligns different data points into a common coordinate system. The image registration process has become an important tool for medical imaging and radiation therapy. This process enables alignment of acquired images of a patient with two different imaging modalities such as MRI and CT. In radiation, treatment planning and delineation, image registration is used to combine the information from e.g. CT and MRI imaging modalities [11]. When image registration is performed, a reference R image and a treatment T image will be defined. By applying image registration process the reference image is kept unaffected and the treatment image is transformed to acquire the spatial and geometry coordinates of reference image.

Image registration can be defined as composing the following components:

- Geometrical transformation, where the treatment image is transformed to reference image.
- Similarity, where it measures how good the registration is performed.
- Regularization, whether the obtained transformation is reasonable.

Each of the above mentioned components are based on which image modality type and registration type is used [11].

2.3.1 Manual Rigid Registration

In the radiation treatment planning process, image registration of OBI and DRR is performed prior to treatment delivery to secure the patient setup and enable correct positioning of the patient. This study involved different imaging modalities such as CT scanner, MRI scanner and On-Board imager. The spatial resolutions of acquired images from the imaging modalities are different, since the images are obtained from different sources and use different imaging devices.

However, a manual rigid registration is performed for patient setup verification at each fraction which transforms an image using translation and rotation. The translation occurs along vertical, longitudinal, lateral and rotation (pitch and rnt).

Figure 2.6 illustrates image manual registration methods, frontal and lateral match, respectively. Frontal matching allows shifts along longitudinal, lateral

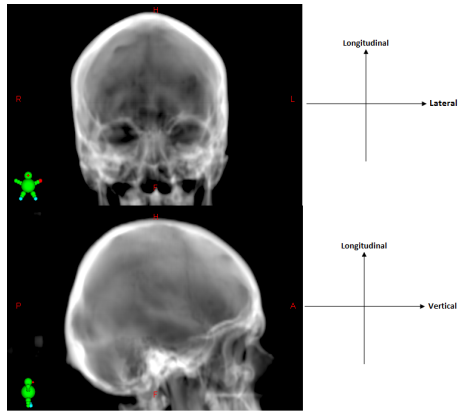


Figure 2.6: The figure at the top illustrates frontal match and the figure at the bottom shows lateral match. The frontal match is carried out by longitudinal, lateral direction and rotation (rnt). The lateral match uses longitudinal, vertical direction and rotation (pitch).

direction and rotation (rnt) and lateral match allows shifts along longitudinal, vertical direction and rotation (pitch).

DRRs are assigned as treatment images and the OBIs are assigned as reference images. The bony structures that appear on a DRR image will be matched with the structures on OBI. The most common structures that are used for matching of palliative patients receiving cranial RT, are shown in figures 2.7 and 2.8. Figures 2.7 shows lateral CT DRR image of a patient receiving whole brain RT. The anatomical structures that are typically used for a lateral match are stated.

Figure 2.8 states the anatomical structures that are typically used for a frontal match of a patient receiving whole brain RT.

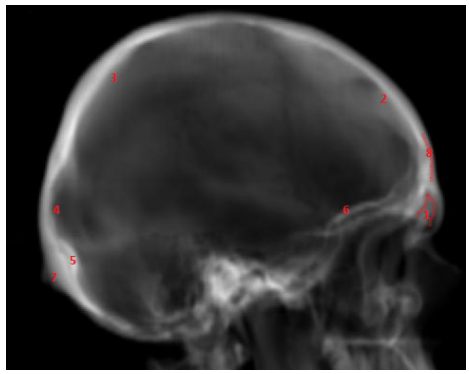


Figure 2.7: The figure illustrates lateral CT generated DRR of a patient receiving whole brain RT. The following anatomical structures are typically used for a lateral match; 1: Sinus frontalis, 2: Os frontale, 3: Os parietale, 4: Os Occipitale, 5: Protuberantia occipitalis intern, 6: Pars Orbitalis ossis sphenoidalis, 7: External occipital protuberance and 8: Lamina externa.

Figure 2.8: The figure illustrates frontal CT DRR image of a patient receiving whole brain RT. The following anatomical structures are used for a frontal match; 1: eyes, 2: Sinus frontalis, 3: Nasal septum.

2.3.2 Affine Transformation

Affine transformation is used for geometrical transformation of an image using translation, rotation, scaling and sharing. In this study the affine transformation is used to find the scaling factor between OBIs and DRRs. Prior to performing image deformation, a linear interpolation is used to evaluate the image intensities at spatial position. The linear interpolation is determined as a weighted sum of the intensities from neighbouring voxels. A 2D linear interpolation is given as follows;

$$I(y) = I(p^1, p^2)(1-\xi^1)(1-\xi^2) + I(p^1+1, p^2)\xi^1 * (1-\xi^2) + I(p^1, p^2+1)(1-\xi^2) * \xi^2 + I(p^1+1, p^2+1)\xi^1 * \xi^2 \quad (2.5)$$

where $I(y)$ is the intensity value which is determined as a weighted sum of the intensities from neighbouring voxels at p^1 p^2 positions. The weights are computed by the zero to one normalized distance ξ^1, ξ^2 to the nearest neighbouring voxels.

To transform x_i coordinate, a affine transformation in 2D is defined as follows [11];

$$y(x_i; A, t) = Ax_i + t \quad (2.6)$$

where t is a translation vector, x is a vector of pixel coordinates and A is a 2x2 matrix describing rotation, translation, scaling and shearing. This can be rewritten as follows;

$$y_i = R*Z*S*x_i+t, R = \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} Z = \begin{bmatrix} 1 & z_x \\ z_y & 1 \end{bmatrix} S = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \quad (2.7)$$

where R is rotation, Z is sharing and S denotes scaling. The model can also be defined as follows;

$$y_i = Q(x_i) * w \quad (2.8)$$

where i defines voxel coordinates and w contains all parameters of an affine model. An optimization algorithm is used to estimate appropriate parameters for the above affine model. The concept of this algorithm is to search the parameters space by adjusting each parameter in turn until no further change appears in the model.

Chapter 3

Methods & Materials

3.1 Data Acquisition

This study includes data from four palliative patients receiving cranial RT. Each patient data consists of planning CT scans, MR scans and anterior and lateral setup (2D) kV radiographs from each fraction. The planning CT scans are acquired with a Philips Brilliance Big Bore CT and the OBIs are acquired at the LINAC with the on-board imager (OBI) at each fraction. The MR scans are obtained with 1 Tesla open MRI-system, Philips Panorama. The MR scans consist of a T1 weighted, DIXON, and UTE sequence data. In this study only UTE sequence MR data is used and the rest are excluded.

Patient	MRI scan	CT scan	Fractionx2 OBI
1	UTE	Planning CT data	9x2
2	UTE	Planning CT data	1x2
3	UTE	Planning CT data	8x2
4	UTE,	Planning CT data	8x2

Table 3.1: The table represents the data from four patients which are divided in CT scan, MR scan and two orthogonal OBIs for each fraction at frontal and lateral position, respectively.

Table 3.1 illustrates data from all four patient receiving whole brain RT with 2D setup verification. The first patient received nine fractions. The second patient received one single fraction RT, while the third and fourth patients each received eight whole brain RT with 2D setup verification.

3.2 Data Processing

The data was anonymized using ConQuest DICOM server 1.4.16. The figure below illustrates data processing in this study.

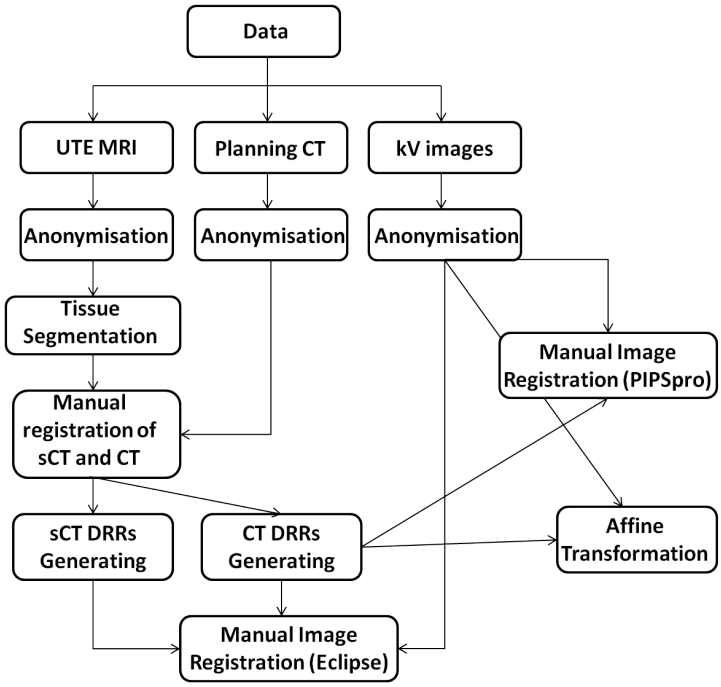


Figure 3.1: The figure illustrates data processing.

3.2.1 Tissue Segmentation

The output from MR UTE sequence was segmented into different tissue classes. MRF was used as a segmentation strategy to classify each data set into air, soft tissue and cortical bone. This was done in MATLAB where an automatic training with 7 classes was shared for MRF with $\beta = 0.7$ and 10 iterations and created so-called substituted CT (sCT) scans [10]. This was performed for all four patients.

3.2.2 DRRs Generating

The patient data sets were imported into Eclipse v.10.0 (Varian Medical Systems) TPS in a training box (T-box). All data were in DICOM format. The planning CT scans and plan for each patient were exported from the clinical system to TPS at the T-box. The plan was transferred to patient planning CT scans. The sCT scans were also imported to TPS at T-box. The plan from planning CT scans were then transferred to sCT scans. The planning CT scans were registered with sCT using TPS manual rigid registration where the planning CT scans were considered as true data as shown in figure 3.2. sCT and CT DRR were generated in TPS and this was done with both anterior and lateral direction, respectively.

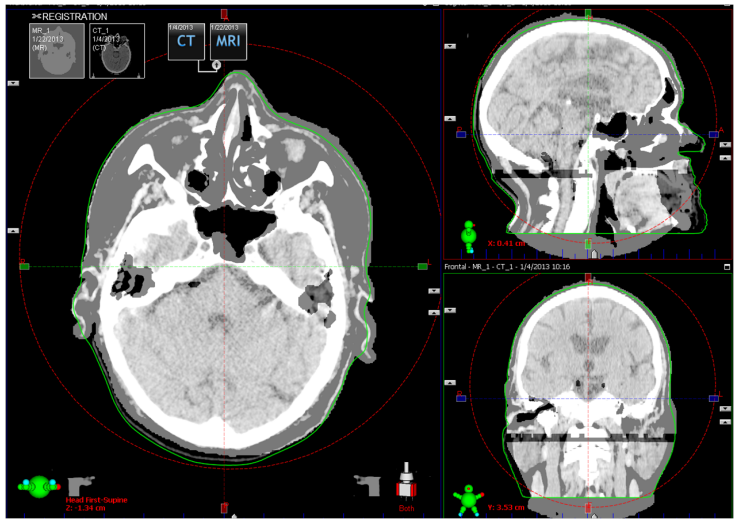


Figure 3.2: Manual rigid registration of CT and sCT scan.

3.2.3 2D Manual Image Registration

In this study it was necessary to find a user-friendly software for RTTs to perform a 2D manual images registration. Initially, the software Portal Image Processing System V.4.2 (PIPSpro) was examined for this purpose. Only one patient data set was applied for this examination. The rest of the 2D image registration was done with Offline Review Eclipse V.10.0 (Varian Medical System).

3.2.3.1 2D Manual Image Registration with PIPSpro

PIPSpro software was used to perform 2D images registration. The application of PIPSpro is described in more detail in section A.1.

sCT DRRs and CT DRRs were used to make a 2D registration in PIPSpro. PIPSpro requires same size images. DRRs and OBIs differ in size, since the images are acquired with three different imaging modalities MRI, CT and OBI, respectively. MATLAB was used to resize DRR to OBI and obtained the same size and resolution, as shown in figure 3.3. Registration in PIPSpro was made

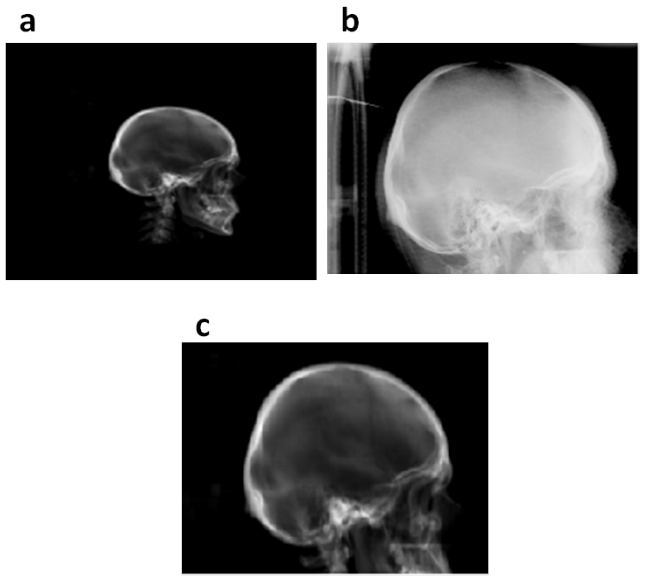


Figure 3.3: Figure a and b illustrates the original CT DRR and OBI, respectively. Figure c shows the resized CT DRR.

by delineating contours around the region of interest (ROI) and drawing landmarks in reference (DRRs) and treatment image (OBIs). The landmarks at reference image were transformed into the treatment image. Figure 3.4 shows the procedure of manual image registration in PIPSpro. Figure 3.4 shows that

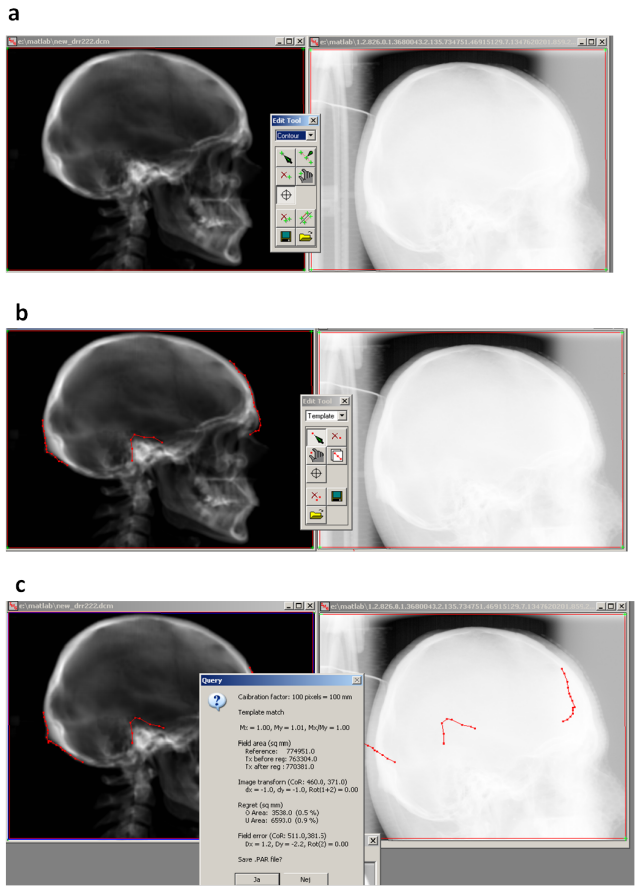


Figure 3.4: Figure illustrates the procedure of manual registration in PIPSpro. a: manual contour delineation, b: manual landmark positioning and c: landmark transformation.

the transformed landmarks from the reference image cannot be aligned with the treatment image, since the treatment images are larger compared to the reference image even though the images were resized to the same size using the information provided by the DICOM-Header.

3.2.3.2 Affine Transformation

As mentioned in the previous section, OBIs and DRRs were not the same size, even though resizing was performed using DICOM information in MATLAB.

Further study was carried out to find the scaling factors between OBI and DRRs. Affine transformation was performed where OBI image was assigned as reference image and DRR as treatment image. The treatment images were transformed and scaled as the reference image and following factors were provided from the affine transformation output: translation, rotation, scaling and shearing. This was done in Matlab. Only one patient data set was used for this study.

3.2.4 2D Manual Image Registration with Eclipse

OBIs from each fraction were imported from the clinical system into TPS at T-box. The relevant OBI images were attached to the relevant DRR images both for CT and sCT DRRs, respectively. This was done for all four patients.

Image registration was performed according to clinical protocols at Herlev Hospital. Offline Review Eclipse v.10 (Varian Medical System) allows both automatic registration and manual registration. The available tools in Offline Review allow numerical shifts along lateral, longitudinal, vertical and two rotations rnt (frontal) and pitch (lateral), together with matching of OBIs and DRRs.

Five experienced RTTs from the clinic were recruited to match OBIs with CT and sCT generated DRRs for each patient in a random order using Manual Match. The sCT and CT generated DRRs were blinded and the RTTs were allowed using of all help functions. The OBIs were assigned as reference images and the DRRs were assigned as treatment images. The RTTs performed matching by aligning the anatomical structures that appear in OBI images with DRRs or vice versa.

As mentioned before, the match of images for each patient were performed in random order, with an RTT match of OBI images first with CT DRRs and than with sCT DRRs. Both lateral and frontal matches were performed in Offline Review by blending these images as shown in figure 3.5.

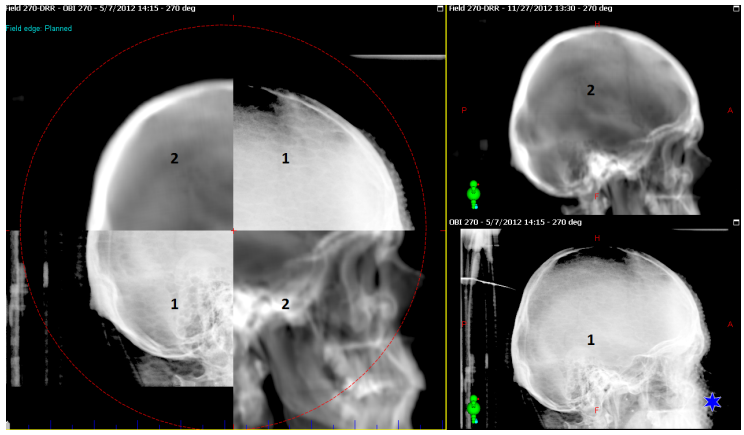


Figure 3.5: The figure illustrates a lateral match where 1 defines OBI image and 2 refers to CT DRR.

3.3 Statistical Approaches

In this study, a statistical analysis was carried out to evaluate whether there is a significant difference in 2D setup verification of patients when performing matching on sCT generated DRRs as compared to CT based DRRs. The match of OBIs with CT and sCT DRRs of each patient were performed in random order. The matched variables are divided into continuous and categorical variables as shown in tables 3.2 and 3.3, respectively.

	Continuous variable
1	Longitudinal (Front)
2	Longitudinal (Lateral)
3	Vertical
4	Lateral
5	rnt
6	Pitch

Table 3.2: The table represents continuous variables.

	Categorical variable
1	Image modality
2	Patient
3	Fraction
4	RTT/Nurse

Table 3.3: The table represents categorical variables.

The statistical software SAS 9.3 was used for all statistical analysis in this study. The significance level is chosen to be $\alpha = 0.05$

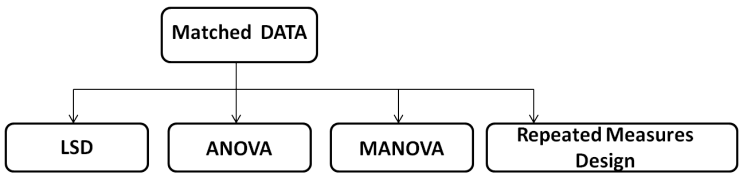


Figure 3.6: The figure illustrates the applied statistical tests on the matched data.

The figure 3.6 illustrates the applied statistical tests on the matched data.

3.3.1 Least significant difference

Fisher's Least significant difference (LSD) test compares all pairs of means with the null hypotheses $H_0 : \mu_i = \mu_k$ against all alternatives (for all $i \neq k$) using t-test [4].

$$t_0 = \frac{\overline{y}_i - \overline{y}_k}{\sqrt{\frac{2MS_E}{n}}} \tag{3.1}$$

where $\overline{y}_i$ and $\overline{y}_k$ are treatment means and MS_E denotes means square error. If the pair of means $\overline{y}_i$ and $\overline{y}_k$ is significant different, H_0 is rejected:

$$|\overline{y}_i - \overline{y}_k| \geq LSD \tag{3.2}$$

where the least significant difference (LSD) is; $LSD = t_{1-\frac{\alpha}{2}}\sqrt{MS_E(\frac{2}{n}}$.

Initially, LSD test based on all continuous variables was carried out. The continuous variables are divided into lateral and frontal match. The lateral match contains the continuous variables, longitudinal (lateral) direction, vertical direction and pitch, respectively. The frontal match contains the continuous variables such as longitudinal (front) direction, lateral direction and rnt.

3.3.2 Analysis of Variance

Analysis of variance (ANOVA) is used to test different factors which change the outcome significantly.

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \tag{3.3}$$

for $i = 1, 2, ..., a$ and $j = 1, 2, ..., n$ In equation 3.3, μ is overall mean, τ_i is a parameter unique to ith treatment called the ith treatment effect and ϵ_{ij} is a random error that appears in the experiment [14]. The equation 3.3 is used to test the null-hypotheses of whether the populations differ significantly or not. The null hypotheses, H_0, states that the means are equal and treatment effects within the different levels, τ_i, replaces by zeros. H_0 is paired with a second, alternative hypotheses, H_1, which means that at least one of i is non-zero. The analyses of variance uses F-test for the hypotheses of no difference in treatment means.

$$F = \frac{\sum_{i=1}^{a}(\overline{Y}_i - \overline{Y})/(a-1)}{\sum_{i=1}^{a}\sum_{j=1}^{n}{}_i(\overline{Y}_{ij} - \overline{Y}_i)^2/(N-a)} = \frac{SS_{Tr}/(a-1)}{SS_E/N-a} \tag{3.4}$$

Y is mean of all measurements, Y_i is measurement of each i and $N = a * n$. The F is distributed with a-1 and N-a degrees of freedom (DF). Table 3.4 illustrates a simple ANOVA table.

Source of Variation	Degrees of freedom	Sum of Squares	Mean Square	F_0
Treatments	a-1	SS_{Tr}	$MS_{Tr} = \frac{SS_{Tr}}{a-1}$	$\frac{MS_{Tr}}{MS_E}$
Error	N-a	SS_E	$MS_E = \frac{SS_E}{N-a}$	
Total	N-1	SS_T		

Table 3.4: The table shows ANOVA table.

An ANOVA test was performed based on only one continuous variable, longitudinal (front) direction. A Reduced model method based on ANOVA was used where all continuous variables were considered. A model was considered with three main factors: modality, RTT and patient, respectively.

3.3.3 Repeated Measures Design

A repeated measures design is appropriate when multiple measures of dependent variables are taken on the same object under different conditions or over two or more time periods. Repeated measures experiments are considered as a factorial design experiment. Responses measured on the same object can be correlated since the responses contain a common contribution from the object. Responses that are measured around the same time can often be correlated better than responses measured far apart in time. The variances of repeated measures change with time. These factors produce a complicated covariance structure of the repeated measures, therefore appropriate statistical analysis is required because of the covariance structure. Univariate and multivariate ANOVA are often performed for the statistical analysis of repeated measures data. In some cases the mixed model methodology is also used for analysis of repeated measures data [14, 13, 24]. A repeated measures design is used in this study. Multiple response is taken in sequence on the same patient. The aim of repeated measures analysis is to examine and compare response trends over time/fraction. This involves comparisons of treatments over fraction and comparison of fraction within a treatment [13]. Univariate and multivariate ANOVA are used for the significant tests.

The univariate ANOVA is valid, if all measurements have equal variance at all fractions and pairs of measurements on the same patient are equally correlated, regardless of the time lag between the measurements. Huynh-Feldt (H-F) and Greenhouse-Geisser (G-G) conditions are necessary for the validity of univariate ANOVA tests. The H-F ϵ and G-G ϵ predict how well the circularity assumption has been met. It ranges from $1/df_{fraction} \leq \epsilon \leq 1$. [13, 24].

The multivariate ANOVA, so-called MANOVA is also used in testing within

subject effects. MANOVA involves four different multivariate tests with different focus. The P-values are obtained from approximate F-test. The four multivariate test statistics are listed below.

- Wilk's Lambda
- Pillai trace
- Hotelling-Lawley trace
- Roy's maximum

The most known test statistic is Wilk's Lambda and is therefore used in study.

Analysis of each continuous variable with 10 repeated measures where fraction was taken as repeated factor SAS generalized linear model (GLM) procedure was used for the implementation of repeated measures method.

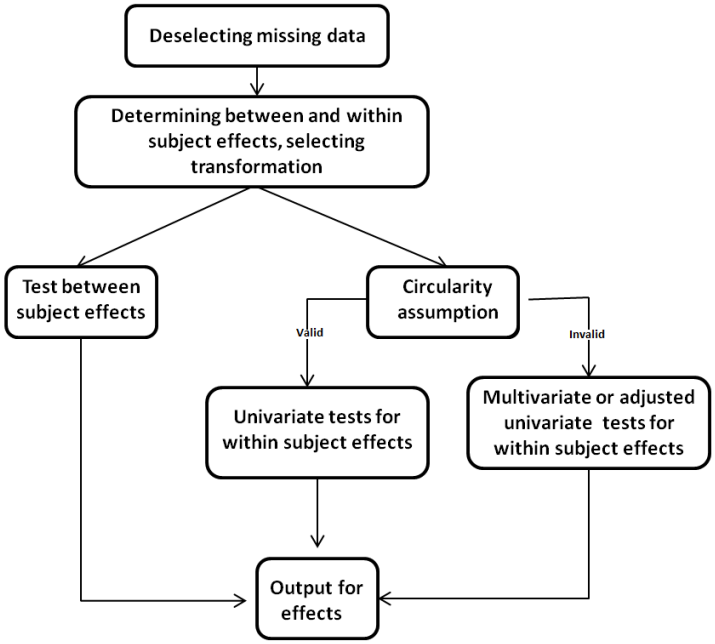


Figure 3.7: The figure illustrates repeated measure method implementation using SAS (PROC GLM).

Figure 3.7 shows implementation of repeated measures strategy in SAS. Initially some of the missing data was deselected. Between subject effects and

within subject effects were determined and transformation was selected. PROC GLM performs a standard significance test between subject effects. The PROC GLM tests whether this structure is significant or not, using the circularity assumption. The mauchly chi-square is used to determine whether the circularity assumption is valid or not, by using H-F condition. If the circularity assumption is valid or significant, then univariate tests within subject effects is used. If the circularity assumption is invalid, then PROC GLM offers two ways to test the significance of within subject effects. 1) to adjust the univariate tests, H-F and G-G adjustments and 2) multivariate tests involving four different tests.

Chapter 4

Results

4.1 MRF Segmentation

Each UTE MRI patient data set was segmented into air, soft tissue and cortical bone using a Markov Random Field classifier, to generate a so-called sCT scan.

An axial slice of the sCT generated scan for all four patients is shown in figure 4.1, where the black color seen in images refers to air, the gray color denotes soft tissue and the white color refers to cortical bone. The figure at the top left shows an axial slice of a sCT scan of patient 1, where the segmented cortical bone appears with a thick contour. This indicates that the cortical bone is over-segmented compared to the soft tissues. The figure at the top right illustrates a sCT slice for patient 2, the cortical bone is segmented when it reaches the occipital part of the cranium, but there is a signal loss at the frontal part. The figure at the bottom left illustrates a sCT slice for patient 3, the cortical bone appears as a bright contour in the image. There are clear artefacts at frontal and occipital parts of the cranium. The bottom right shows an axial sCT slice of patient 4, where the cortical bone is segmented. All patients suffer from artefacts, especially at frontal and occipital parts of the cranium, which are very clear in the images.

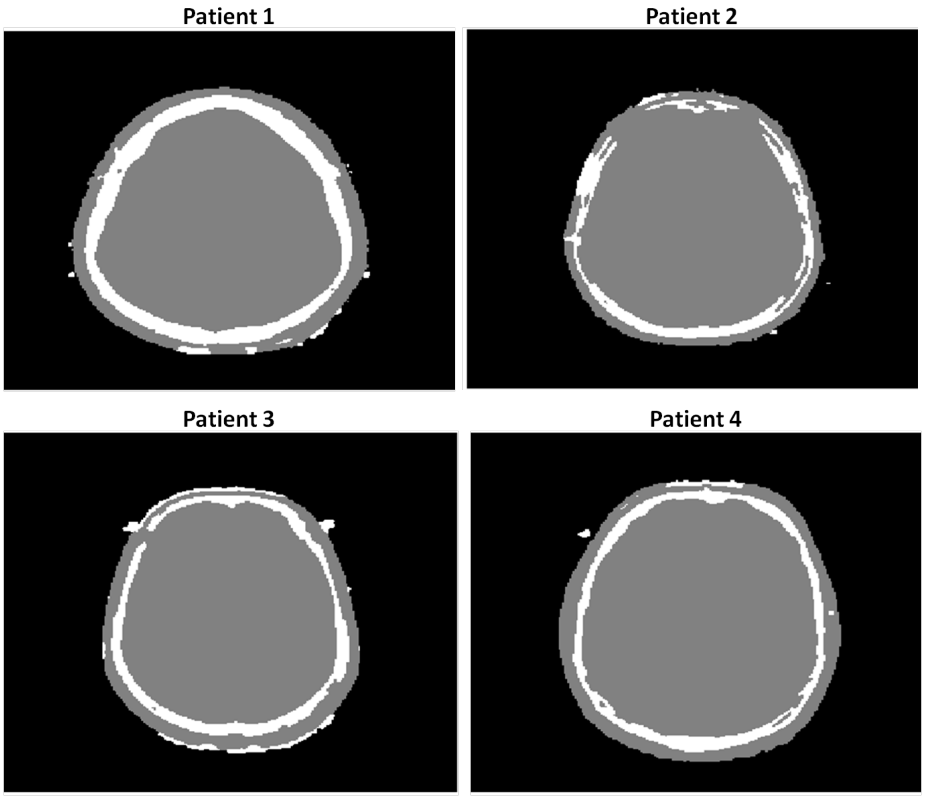


Figure 4.1: The figure illustrates axial slice of sCT scan for all four patients, patient 1, patient 2, patient 3 and patient 4, respectively. The black color seen in the images refers to air, the gray color denotes soft tissue and the white color refers to cortical bone.

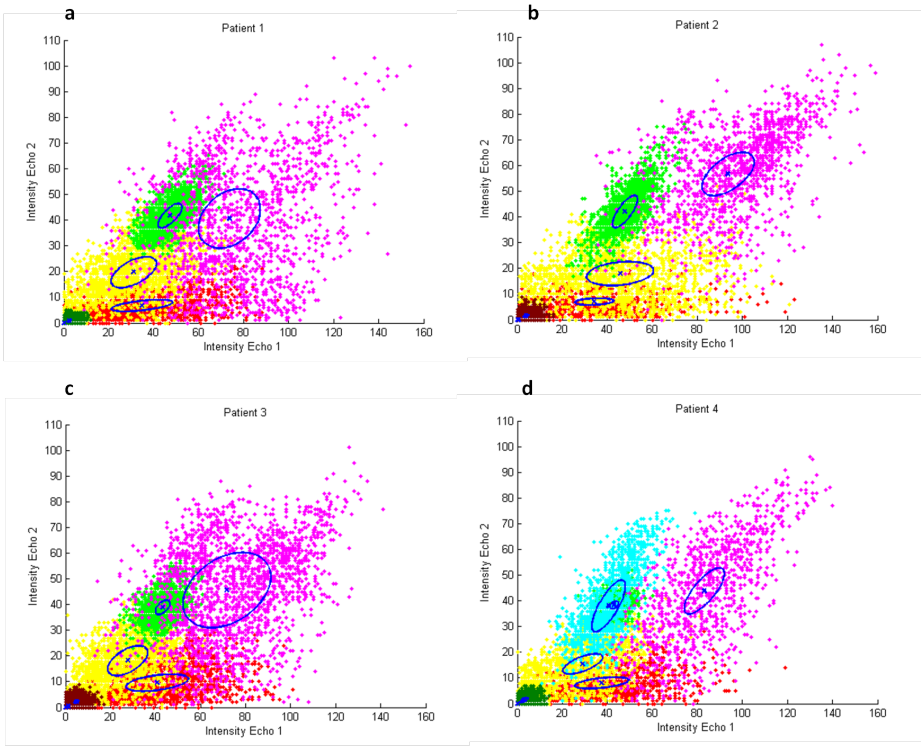


Figure 4.2: The figure illustrates intensity plots of patient 1, patient 2, patient 3 and patient 4, respectively. Each coloured region refers to a tissue group, where dark red and dark green refer to air, red denotes bone, yellow, green, purple and light blue refer to different soft tissue groups. The x-axis refers to intensity echo 1 and the y-axis is intensity echo 2.

Figure 4.2 illustrates intensity plots for all four MRF segmented patient data. The tissue groups or classes are shown in different colors. The dark green and dark red color refer to air, red denotes bone, yellow, green, purple and light blue refer to different soft tissue groups. Interesting observations are made when looking at the intensity plot for patient 1 and patient 3, where the purple tissue group covers two Gaussian distributions. This indicates a poor classification of the soft tissue for patient 1 and 2. A better classification could have been obtained if the MRF classification algorithm was performed until a satisfying result was obtained. To avoid bias the MRF classification algorithm was performed only once.

4.2 Comparison of sCT and CT generated DRRs

The geometrical evaluation of all four patients was done by using the dice coefficient, which measures the similarity of sCT bone volume and CT bone volume. The dice coefficient (D) is calculated by the given intersection volume $A \cap B$ and the individual volumes A and B [3];

$$D = \frac{2(A \cap B)}{A + B} \tag{4.1}$$

The perfect overlap of A and B volumes acquires $D = 1$ whereas two disjoint volumes lead to $D = 0$. A refers to CT bone volume and B refers to sCT

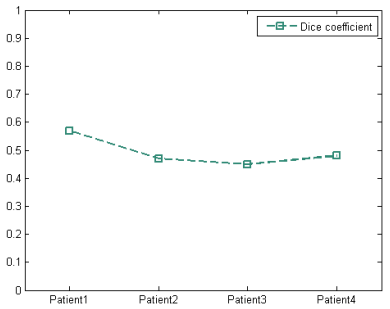


Figure 4.3: The figure illustrates the determined dice coefficients for all four patients.

bone volume. The dice coefficients for all four patients were calculated from the equation 4.1 and are plotted in figure 4.3. The mean and standard deviation of dice coefficient was determined to be 0.50 ± 0.05.

The bone sCT DRRs were generated from sCT scan and the bone CT DRRs were generated from planning CT scan. This was done both for lateral and frontal direction.

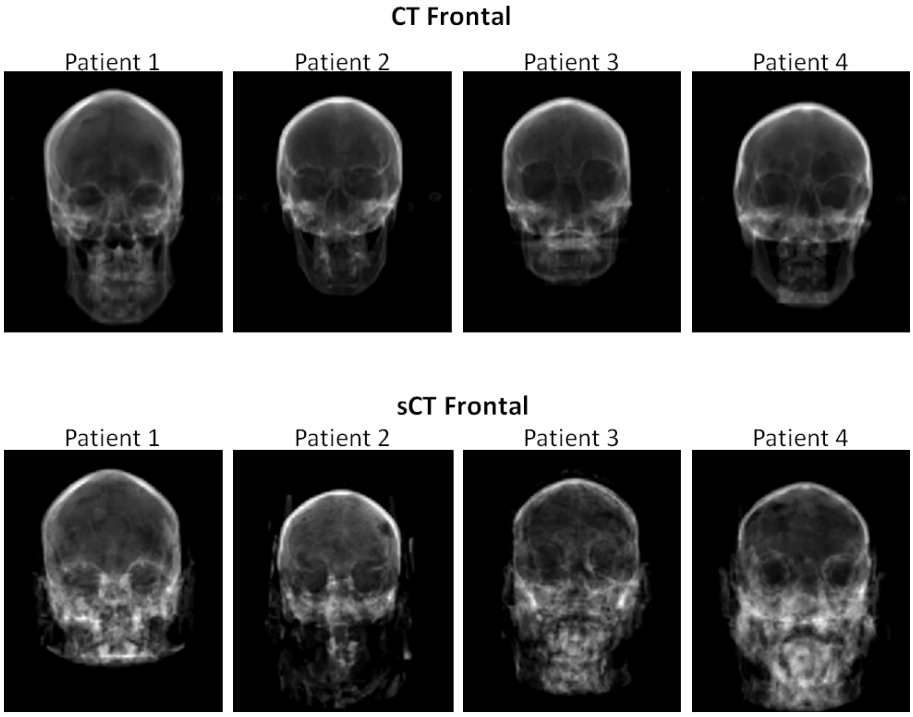


Figure 4.4: The figure illustrates CT and sCT generated DRRs images for the frontal directions for patient 1, patient 2, patient 3 and patient 4, respectively.

Figure 4.4 shows frontal CT and sCT generated DRRs for patient 1, patient 2, patient 3 and patient 4, respectively. The sCT DRRs for patient 1, 2 and 4 are close to the CT DRRs when looking at the eye part of the patients. Patient 3 has obtained a poor bone segmentation when looking at the eye part and performs with the less dice coefficient shown in figure 4.3. The bone segmentation around the cranium for all 4 patients is close to the CT DRRs shown in figure 4.4.

Lateral CT and sCT DRRs are shown in figure 4.5 for patient 1, patient 2, patient 3 and patient 4. The sCT DRRs for patient 1 and 4 are close to the CT DRRs when comparing the bone structures, especially the bone structure along pars orbitalis ossis sphenoidalis (shown in figure 2.7 for lateral match in chapter 2). Patient 1 and 4 perform with better dice coefficient compared to patient 2

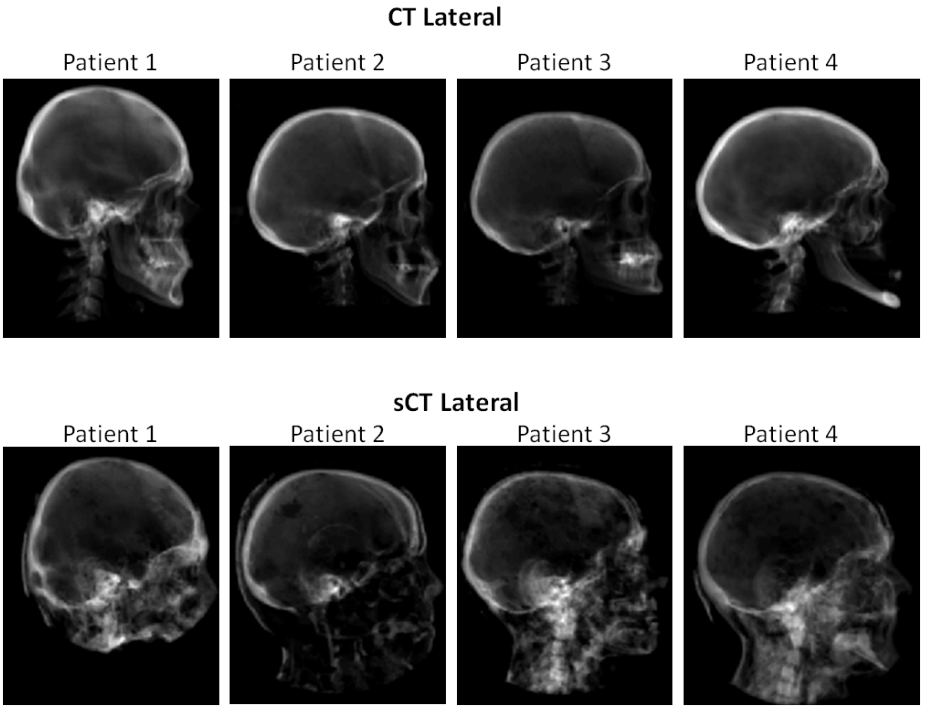


Figure 4.5: The figure shows CT and sCT generated DRRs images for the lateral directions for patient 1, patient 2, patient 3 and patient 4, respectively.

and 3 shown in figure 4.3. In addition, the bone segmentation along external and internal occipital part is close to the CT DRRs for all patients shown in figure 4.5.

4.3 Affine Transformation

Affine transformation was performed to find the scaling factors between OBI and DRRs, where OBI image was assigned as a reference image and DRR as treatment image. The following parameters were provided from the affine transformation output: translation, rotation, scaling and shearing.

Fraction	T_1	T_2	R	s_x	s_y	z_x	z_y
1	-6.433	-3.878	0.072	0.676	0.630	-0.050	0.049
2	-7.776	3.297	-0.098	0.683	0.606	0.136	-0.177
3	-4.948	-1.524	-0.032	0.641	0.643	0.043	-0.070
4	-5.573	-1.372	0.011	0.647	0.646	0.008	-0.035
5	-6.713	-1.731	0.264	0.669	0.612	-0.207	0.249
6	-8.330	-0.292	0.156	0.687	0.624	-0.079	0.096
7	-8.181	-0.568	0.149	0.687	0.623	-0.080	0.097
8	-6.569	-2.120	0.093	0.645	0.636	-0.072	0.037
9	-6.773	-1.368	0.045	0.647	0.642	-0.019	-0.012
10	-6.976	-2.517	0.088	0.646	0.638	-0.069	0.038

Table 4.1: The table represents the output parameters of affine model after final estimate. T_1 and T_2 are the translation parameters, R is the rotation parameter given radian, s_x and s_y are the scaling parameters and z_x and z_y are the shearing parameters.

Table 4.1 illustrates the output parameters of the affine model. T_1 and T_2 are the translation parameters, R is the rotation parameter given radian, s_x and s_y are the scaling parameters and z_x and z_y are the shearing parameters. For fraction 1, 2, ..,7 appropriate affine model parameters were obtained with $s_x = 0$ and $s_y = 0$ initial estimated value, while for fraction 8, 9 and 10 three estimated values were necessary until appropriate affine model parameters were obtained.

Figure 4.6 illustrates affine transformation process for fraction 10. The figure at the top left shows CT DRR, the figure top right illustrates OBI and the image at the bottom shows the output of transformation with a final estimate $s_x = 0.679$ and $s_y = 0.630$.

The scaling factors s_x and s_y were varied for each fraction and in some cases as for fraction 8 and 10 more than one estimated value was needed.

This study indicates an error in resolution value (mm/pixel), which is provided in DICOM header for OBI and DRRs. The provided resolution values for all OBIs are fixed and are given [0.388 0.388]mm/pixel. The resolution values for

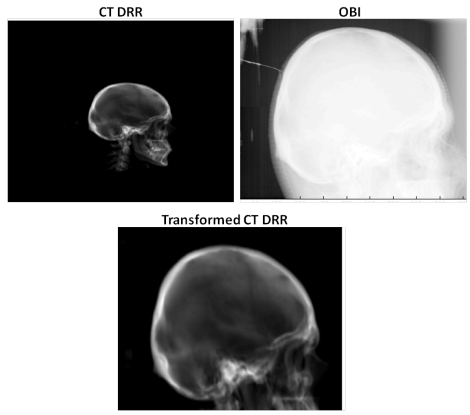


Figure 4.6: The figures illustrates the affine transformation process for fraction 10. The figure at the top left illustrates CT DRR, the figure top right shows OBI and the image at the bottom shows the output of transformation with a final estimate $s_x = 0.679$ and $s_y = 0.630$.

CT DRRs are also fixed and [0.976 0.976]mm/pixel. These DICOM header provided resolution values concern OBIs and CT DRRs for all four patients. Therefore use of PIPSpro software for manual image registration was not appropriate for this study.

4.4 CT and sCT generated DRR Match

Five RTTs from the clinic enabled performance of manual matching of OBIs with CT and sCT DRRs, respectively. RTTs performed both lateral and frontal match. This results in 260×6 matched data points which can be used for statistical evaluation if there is a significant difference between matches performed on CT DRRs and sCT DRRs in 2D patient setup verification.

4.5 Statistical Analysis

4.5.1 LSD

The LSD test for all continuous variables (shifting directions) was calculated to distinguish between lateral match and frontal match when matching with sCT and CT generated DRRs. The significance level is chosen at $\alpha = 0.05$

Shifting Directions	$\overline{y}_{MR}$	$\overline{y}_{CT}$	MS_E	$\overline{y}_{sCT} - \overline{y}_{CT} \geq LSD$
Longitudinal (Front)	-0.0346	-0.0307	0.0519	-0.0038<0.0556
Longitudinal (Lateral)	0.0477	0.0085	0.0304	0.0392<0.0426
Vertical	0.0577	0.0131	0.0187	0.0446>0.0334
Lateral	-0.0577	-0.0684	0.0276	0.0107<0.0406
Pitch	0.4177	0.7477	1.1380	-0.3300<0.2606
rnt	-0.1667	-0.4569	1.7820	0.2886<0.3261

Table 4.2: The table illustrates comparisons among the observed modality averages, where mean MR $\overline{y}_{MR}$ and mean CT $\overline{y}_{CT}$ are in mm, MS_E denotes mean square error and $\overline{y}_{MR} - \overline{y}_{CT}$ is the difference between mean sCT and CT and LSD is the determined least significant difference.

The LSD analysis shown in table 4.3, shows a non-significant difference between mean CT and sCT except vertical direction ($0.0446 > 0.0334$). There is a significant difference in sCT and CT DRRs in vertical direction when performing lateral match. This may cause the poor segmentation sCT along pars orbitalis ossis sphenoidalis for patient 2. Therefore the matched data for patient 2 was excluded and a new LSD test based on matched data for patient 1, 3 and 4 was performed.

The LSD test still shows a significant difference between sCT and CT for vertical direction when performing lateral match. By excluding patient 2 the LSD result is not affect for the vertical directions.

Shifting Directions	$\overline{y}_{MR}$	$\overline{y}_{CT}$	MS_E	$\overline{y}_{sCT} - \overline{y}_{CT} \geq LSD$
Longitudinal (Front)	-0.0360	-0.0320	0.0539	-0.0040<0.0574
Longitudinal (Lateral)	0.0536	0.0012	0.0307	0.0416<0.0437
Vertical	0.0512	0.0136	0.0170	0.0376>0.0325
Lateral	-0.0568	-0.0664	0.0285	0.0096<0.0421
Pitch	0.4208	0.7776	1.1627	-0.3550<0.2680
rnt	-0.1744	-0.4889	1.8353	0.3152<0.3375

Table 4.3: The table illustrates comparisons among the observed modality averages, where mean MR $\overline{y}_{MR}$ and mean CT $\overline{y}_{CT}$ are in mm, MS_E denotes mean square error and $\overline{y}_{MR} - \overline{y}_{CT}$ is the difference between mean sCT and CT and LSD is the determined least significant difference. Patient 2 is deselected from matched data.

4.5.2 ANOVA

An ANOVA test is used to measure the variation of means among modalities, patients, fractions and RTTs, respectively. A fixed model is considered where the factors such as modality, patient fraction and RTT (RTT donates nurse) are deterministic. The null-hypothesis, H_0 for sources such as modality, patient, fraction and RTT and the interaction between, is that all the means, μ for the source, are the same and the alternative hypothesis, H_1 states at least one of the means differs from the others.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	259	13.39211538	0.05170701	.	.
Error	0	0.00000000	.		
Corrected Total	259	13.39211538			

R-Square	Coeff Var	Root MSE	LongF Mean
1.000000	.	.	-0.032692

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Modality	1	0.00096154	0.00096154	.	.
Nurse	4	0.39423077	0.09855769	.	.
Modality*Nurse	4	0.17192308	0.04298077	.	.
Patient	3	0.31936538	0.10645513	.	.
Modality*Patient	3	0.03428846	0.01142949	.	.
Nurse*Patient	12	1.12796368	0.09399697	.	.
Modali*Nurse*Patient	12	0.39949359	0.03329113	.	.
Fraction	9	0.73841667	0.08204630	.	.
Modality*Fraction	9	0.37609524	0.04178836	.	.
Fraction*Nurse	36	1.53157341	0.04254371	.	.
Modali*Fractio*Nurse	36	0.95554167	0.02654282	.	.
Fraction*Patient	13	2.34933333	0.18071795	.	.
Modali*Fracti*Patien	13	0.79765476	0.06135806	.	.
Fracti*Nurse*Patient	52	2.33623214	0.04492754	.	.
Moda*Frac*Nurs*Patie	52	1.85904167	0.03575080	.	.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.00456026	0.00456026	.	.
Nurse	4	0.23526518	0.05881630	.	.
Modality*Nurse	4	0.15687969	0.03921992	.	.
Patient	3	0.26541667	0.08847222	.	.
Modality*Patient	3	0.04042857	0.01347619	.	.
Nurse*Patient	12	1.12526786	0.09377232	.	.
Modali*Nurse*Patient	12	0.41579167	0.03464931	.	.
Fraction	9	0.73841667	0.08204630	.	.

Figure 4.7: The figure shows ANOVA tables for the longitudinal (front) direction.

Figure 4.7 illustrates ANOVA tables for longitudinal (front) direction. The f-values and P-values are not determined, since there is no error for degrees of freedom.

A new model is considered based on three factors, modality, patient and RTT. An ANOVA test is performed to test the variation of means among modalities, patients and RTTs, respectively.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	39	2.44822650	0.06277504	1.26	0.1521
Error	220	10.94388889	0.04974495		
Corrected Total	259	13.39211538			

R-Square	Coeff Var	Root MSE	LongF Mean
0.182811	-682.2270	0.223036	-0.032692

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Modality	1	0.00096154	0.00096154	0.02	0.8896
Nurse	4	0.39423077	0.09855769	1.98	0.0984
Modality*Nurse	4	0.17192308	0.04298077	0.86	0.4863
Patient	3	0.31936538	0.10645513	2.14	0.0961
Modality*Patient	3	0.03428846	0.01142949	0.23	0.8756
Nurse*Patient	12	1.12796368	0.09399697	1.89	0.0368
Modali*Nurse*Patient	12	0.39949359	0.03329113	0.67	0.7800

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.00032653	0.00032653	0.01	0.9355
Nurse	4	0.11616128	0.02904032	0.58	0.6747
Modality*Nurse	4	0.05140391	0.01285098	0.26	0.9044
Patient	3	0.31936538	0.10645513	2.14	0.0961
Modality*Patient	3	0.03428846	0.01142949	0.23	0.8756
Nurse*Patient	12	1.12796368	0.09399697	1.89	0.0368
Modali*Nurse*Patient	12	0.39949359	0.03329113	0.67	0.7800

Figure 4.8: The figure shows ANOVA tables for the longitudinal (front) direction where the factor fraction is not considered in the model.

Figure 4.8 illustrates ANOVA tables for longitudinal (front) where fraction is not included in the model. By comparing ANOVA tables from the first model (fraction factor included) with the second model (fraction factor not included), it can be seen that the error term in the second model corresponds to the sum of the square of all effects containing fraction factor in the first model. If there is a strong correlation between different fraction for the same patient, this may corrupt the estimate of error variance and may cause bias. A positive correlation can cause a smaller variance. This will again lead to the F-test values increasing and therefore seem more significant. If factor has been found not to be significant, then this is most likely the correct conclusion. A model which takes the eventual correlation between the fractions into account is a repeated measures design. The analysis of repeated model design will be introduced later in this chapter.

4.5.2.1 Reduced Model Method with ANOVA

A model is considered to have three main effects: modality, nurse (nurse refers to RTT), patient and the interactions between them. This model is so-called a full model. An ANOVA test was performed on the full model and the reduced model method was carried out. The full model was reduced by taking out the most non-significant interaction term, i.e. interaction term with highest P-value. This procedure was repeated until a possible significant model was obtained [4]. This was performed for the lateral and the frontal match variables. The lateral match variables are longitudinal (lateral) direction, vertical direction and pitch, respectively. The frontal match variables are longitudinal (front) direction, lateral direction and rnt.

The null-hypothesis, H_0 for sources such as modality, patient and RTTs, is that all the means, μ for the source are the same and the alternative hypothesis, H_1 states at least one of the means differs from the others. For the ANOVA table analysis only the main factors will be analysed.

a

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.10003846	0.10003846	4.10	0.0438
Nurse	4	0.14946154	0.03736538	1.53	0.1931
Patient	3	1.59794124	0.53264708	21.85	<.0001

b

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	7.07850000	7.07850000	7.24	0.0076
Nurse	4	27.38500000	6.84625000	7.00	<.0001
Patient	3	20.75324038	6.91774679	7.07	0.0001

c

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.21371655	0.21371655	17.49	<.0001
Nurse	4	0.12340136	0.03085034	2.52	0.0418
Patient	3	0.42971154	0.14323718	11.72	<.0001
Modality*Patient	3	0.39164316	0.13054772	10.68	<.0001
Nurse*Patient	12	0.94137821	0.07844818	6.42	<.0001
Modali*Nurse*Patient	16	0.34119444	0.02132465	1.74	0.0403

Figure 4.9: The figure illustrates ANOVA tables for reduced model of the lateral match variables. a: longitudinal (lateral) direction, b: pitch and c: vertical direction. The significant factors are marked in red.

Figure 4.9 illustrates ANOVA tables for the reduced model of the lateral match variables. Table *a* shows ANOVA table for longitudinal (lateral) direction, table *b* illustrates ANOVA table for pitch and table *c*shows ANOVA table for vertical direction. A significant difference between CT and sCT was observed for longitudinal (lateral) direction, pitch and vertical direction. The modality difference for vertical direction is more significant compared to the pitch and longitudinal (lateral). There is also a significant difference between nurses when

performing lateral match, except for the longitudinal (lateral) direction. The most significant difference seen in pitch compared to the vertical direction. In the clinic the RTTs do not use pitch and rnt when matching due to patient setup verification. This can be one of the reasons that the RTTs differs significantly in pitch (rotation) when performing lateral match. The ANOVA tables, *a*, *b* and *c* show a significant difference between patients, since the sCT and CT generated DRRs from patient to patient shown in figures 4.4 and 4.5.

a

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.00096154	0.00096154	0.02	0.8904
Nurse	4	0.39423077	0.09855769	1.95	0.1025
Patient	3	0.31936538	0.10645513	2.11	0.0998

b

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	5.4375385	5.4375385	4.11	0.0437
Nurse	4	13.6279231	3.4069808	2.57	0.0383
Patient	3	113.8977404	37.9659135	28.68	<.0001

c

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.00753846	0.00753846	0.35	0.5542
Nurse	4	0.22669231	0.05667308	2.64	0.0346
Patient	3	1.51778846	0.50592949	23.54	<.0001

Figure 4.10: The figure illustrates ANOVA tables for reduced model of the frontal match variables. a:longitudinal (front) direction, b: rnt and c: lateral direction. The significant factors are marked in red.

Figure 4.10 shows ANOVA tables for the reduced model of the frontal match variables. Table *a* illustrates ANOVA table for longitudinal (front) direction, table *b* refers to ANOVA table for rnt and table *c* shows ANOVA table for lateral direction. A significant difference between modality means, CT and sCT shown in rnt. The nurse effect is also observed to be significant in rnt and longitudinal direction.

The modality difference is higher for the lateral direction compared to the frontal direction. Vertical direction has the highest P-value compared to the rest of shifting direction: longitudinal (lateral) direction, longitudinal (front) direction, lateral direction, pitch and rnt, respectively. This can be compared with the LSD test outcome, which showed that the vertical direction was significant.

4.5.3 MANOVA

A MANOVA test was performed where all the continuous variables were considered as responses, such as longitudinal (lateral) direction, longitudinal (front) direction, vertical direction, lateral direction, pitch and rnt, respectively. All

four factors (modality, patient, RTT and fraction) were included in the model, RTT denotes nurse. The purpose of this MANOVA was to test the factor effects when having multivariate responses.

Only Wilk's lambda estimated F-test was considered for the analysis of the factors. The level of significance is chosen to be $\alpha = 0.05$.

a

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Modality Effect
H = Type III SSCP Matrix for Modality
E = Error SSCP Matrix

S=1 M=2 N=118.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.93611990	2.72	6	239	0.0143

b

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Fraction Effect
H = Type III SSCP Matrix for Fraction
E = Error SSCP Matrix

S=6 M=1 N=118.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.58394844	2.52	54	1223.3	<.0001

c

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Nurse Effect
H = Type III SSCP Matrix for Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=118.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.76559097	2.77	24	834.98	<.0001

d

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Patient Effect
H = Type III SSCP Matrix for Patient
E = Error SSCP Matrix

S=1 M=2 N=118.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.60269731	26.26	6	239	<.0001

Figure 4.11: The figure illustrates the MANOVA test hypothesis. a: Hypothesis for modality effect, b: Hypothesis for fraction effect, c: Hypothesis for nurse effect and d: Hypothesis for patient effect.

Figure 4.11 shows the MANOVA test hypothesis for modality, fraction, nurse and patient effects. A significant modality, nurse patient and fraction effects were observed when performing MONOVA test. It is observed that the modality effect is significant at $\alpha = 0.05$ significance level whereas nurse, patient and fraction are significant less $\alpha = 0.01$ significance level. This signifies that modality is less significant compared to the other factors, such as RTT, patient and fraction.

A new MANOVA test was performed where only the data from the first fraction for all continuous variables were considered. Three factors were considered: modality, patient and nurse.

Figure 4.12 shows the MANOVA test hypothesis for modality effect, a, for RTT/nurse effect, b and for patient effect, c, respectively. A significant difference was observed in patients when taking only the first fraction for all four patients. The first fraction was considered because patient 2 had received only one fraction RT. No significant difference was shown between CT and sCT in this analysis. There was also no significant difference between RTTs when performing lateral and frontal match.

a

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Modality Effect
H = Type III SSCP Matrix for Modality
E = Error SSCP Matrix

S=1 M=2 N=13

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.72397764	1.78	6	28	0.1398

b

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Nurse Effect
H = Type III SSCP Matrix for Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=13

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.51239104	0.87	24	98.89	0.6395

c

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Patient Effect
H = Type III SSCP Matrix for Patient
E = Error SSCP Matrix

S=1 M=2 N=13

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.27120729	12.54	6	28	<.0001

Figure 4.12: The figure illustrates the MANOVA test hypothesis. a: Hypothesis for modality effect, b: Hypothesis for nurse effect and c: Hypothesis for patient effect.

4.5.4 Repeated Measure Design

Further analysis was made to include the fraction factor. The analysis of each continuous variable for lateral and frontal match with 10 repeated measures was performed taking fraction as the repeated factor. The fractions were assumed to be correlated, since patient artefacts will most likely influence all fractions and this will cause correlation between fraction. A simple ANOVA, univariate ANOVA and MANOVA were carried out for the analysis of repeated measures data. The difference between a simple ANOVA and a univariate ANOVA is that the univariate ANOVA corresponds to a pooled value of fraction. The significant level is chosen to be $\alpha = 0.05$.

The results from the ANOVA, univariate ANOVA and MANOVA tables are shown and described in detail in section A.2.1. In this section the important test values effects will be pointed out, such as modality, RTT and fraction, respectively.

For the lateral match, the structures for longitudinal (lateral) direction, pitch and vertical direction were valid, since H-F and G-G estimates of ϵ were lower than 1. The simple ANOVA tests, so-called tests of hypotheses for between subjects, showed that the main effect modality and nurses were non-significant for longitudinal (lateral) direction, vertical direction and pitch. This indicates that there is no difference between CT and sCT DRRs and between RTTs when applying repeated measure design analysis. A significant difference in fraction effect was observed for vertical direction and pitch both for univariate ANOVA and MANOVA.

For the frontal match the structures for longitudinal (front) direction, lateral direction and rnt were valid, since the H-F and G-G ϵ satisfy the circularity

assumptions. No differences were observed in means for modality and RTT for longitudinal (front) direction, lateral direction and rnt when performing the ANOVA test. The univariate test showed a significant difference in fraction for lateral direction and rnt. For the MANOVA, test a significant difference was observed for longitudinal (front) and lateral directions.

In general the repeated measures design analysis with both univariate ANOVA and MANOVA tests indicate that there is a significant difference in fraction for both lateral and frontal match variables. No significant difference was observed for the modality, which means no significant difference between CT and sCT DRRs. The difference between the RTTs were also non-significant. This indicates that RTTs are representative when performing repeated measures design analysis.

Chapter 5

Discussion

5.1 MRF Segmentation

The MRF classification strategy facilitates segmentation of the UTE MRI output into air, soft tissue and cortical bone and creates a sCT scan for all four patients. An axial slice for all four patients is shown in figure 4.1 in chapter 4 and all four patients suffer from artefacts. These artefacts can be caused by susceptibility artefact which can appear in air-to-soft tissue boundaries and signal loss. The signal loss can be due to reduced covering of coil since the immobilization devices in some cases can prevent the use of an optimal coil. The artefacts were also apparent both in frontal and lateral sCT DRRs shown in figures 4.4 and 4.5 in chapter 4. The appeared artefacts in the sCT generated DRRs do not disturb the matching procedure since the RTTs aligned the bone structures of OBIs with sCT generated DRRs.

The distribution of different tissue groups or classes are shown as intensity plots in figure 4.2. The tissue classes for patient 2 and 4 are finely distributed compared to the intensity plots for patient 1 and patient 3. The intensity plots for patient 1 and patient 2 that illustrate a soft tissue group covers two Gaussian distributions which can indicate that soft tissue classification for patient 1 and 3 is not as good as patient 2 and 4. The classification of the soft tissue could have been better, but the bone segmentation of all four patients is sufficient to

perform matching.

5.2 Comparison of sCT and CT generated DRRs

The anatomical bone structures of CT and sCT DRRs for patient 1 in figure 4.5 and 4.4 in chapter 4 are comparable, but the bone segmentation at the external occipital part is poor compared to the other patients lateral sCT generated DRRs shown in figure 4.5. The similarity between sCT and CT bone volume was determined to be 0.57 shown in figure 4.3 in chapter 4, which is slightly better compared to the rest of the patients. Since the overall bone segmentation for patient 1 has been performed better compared to the other patients why patient 1 obtained the highest dice coefficient.

Patient 2 has obtained a poor bone segmentation along pars orbitalis ossis sphenoidalis (structure shown in figure 2.7 in chapter 2) compared to the other patients when looking at the sCT generated DRRs. The dice coefficient was determined to be 0.47. This can be due to poor bone segmentation. A better segmentation could have been obtained if the MRF classification algorithm performed until a satisfied segmentation was obtained.

Patient 3 performs with a less determined dice coefficient 0.45 shown in figure 4.3 in chapter 4 compared to the rest of the patients. The bone segmentation at the eye part is very poor and not comparable to the frontal CT generated DRR shown in figure 4.4, but segmented bone around the cranium is comparable to CT generated DRR. This is due to the bone around the cranium is more dense. The lateral sCT and CT are very similar to each other except for the bone structure along the pars orbitalis ossis sphenoidalis.

The similarity between CT and sCT bone volume for patient 4 was determined to be 0.48, which is the next best determined dice coefficient shown in figure 4.3. The sCT generated DRRs bone structures are very close to the CT DRRs bone structures shown in figures 4.5 and 4.4.

5.3 CT and sCT generated DRR Match

In general the RTTs enabled performance of 2D matching of OBIs with CT and sCT generated DRRs. The RTTs found the lateral match much easier compared to the frontal match when performing 2D matching of OBIs with sCT DRRs.

This is due to the bone structures in lateral sCT generated DRRs are more visible compared to the frontal sCT generated DRRs.

On the daily basis the RTTs are used to match the OBI with CT DRRs, they are not familiar with sCT DRRs. The RTTs are therefore more confident when matching OBI with CT DRRs. This influences the outcome of their matching procedure. Even though the sCT and CT DRRs were blinded, the RTTs were still able to differentiate between CT and sCT DRRs. There is also a variation within RTTs when performing matching, i.e. knowledge and experience. This factor affects the matched outcome as well.

Different statistical approaches were carried out to determine the uncertainty in the outcome of the matching due to a difference in modalities, such as CT and sCT modalities and therefore different conclusions were obtained.

Least significant difference (LSD) fisher's test was performed because it is a simple method taking only modality into account to distinguish between lateral and frontal match variables. For the lateral match solely vertical direction was significant whereas longitudinal (lateral) and pitch were non-significant. This means that the sCT and CT modality difference shows in vertical direction. The frontal match variables, longitudinal (front) direction, lateral direction and rnt were non-significant. The differences between the frontal sCT and CT generated DRRs seen to be larger compared to lateral sCT and CT generated DRRs due to bone structures shown in figures 4.5 and 4.4 in chapter 4. The RTT matched bone structures as desired. The RTT carried out matching by aligning the bone structures of sCT and CT generated DRRs that are most visible and comparable to OBI. The frontal match performed mainly focusing on the bone structures along the cranium due to a poor bone segmentation around the eye part when looking at the frontal sCT generated DRRs 4.4. The lateral match performed by aligning more bone structures compared to the frontal match. This can be one of the reasons that vertical direction in lateral match is significant.

The reduced model analysis showed a significant difference between CT and sCT modalities for the lateral match variables, longitudinal (lateral) direction, vertical direction and pitch, respectively. For the frontal only rnt showed a significant difference between sCT and CT modalities p-value 0.0437, which is very close to the significance level $\alpha = 0.05$. In general the reduced model analysis showed that only vertical direction was most significant. This can be compared to LSD test where the vertical direction also was found to be significant.

MANOVA test showed a significance difference between CT and sCT modality when taking all fractions into account. No significant difference was observed when taking only one fraction. This indicates that the fractions provide a small

uncertainty measurement in the experiment.

The repeated measures design analysis with ANOVA test showed no significant difference between CT and sCT modality. The differences between RTTs were also observed non-significant which indicates that RTTs are representative. A significant difference between fractions was obtained when performing univariate ANOVA and MANOVA.

Based on the results from the different statistical approaches it chosen to focus on reduced model analysis, since it is more accurate because the model considers the modality, RTT and patient factors which is known to have influence on the matched outcome.

Chapter 6

Conclusion

The investigation shows that 2D patient setup verification solely based on MRI is a feasible alternative to current CT based RT.

The Markov random field classification strategy was able to segment bone for all four patients. In general the bone segmentation around the cranium both for the lateral and frontal DRRs was sufficient, however the bone segmentation for eye part was poor for the frontal DRRs.

The maximum dice coefficient obtained was 0.57 which is far from the ideal dice coefficient of 1. The result confirms the poor bone segmentation.

For longitudinal (both lateral and frontal) the RTTs were observed non-significant. The rnt and pitch shows a significant difference between RTTs, this confirms lake of experience in rnt and pitch. The lateral and vertical directions show a significant difference between the RTTs, this is due to poor segmentation of fine bone structures, which confirms again disadvantages of MRF classification strategy.

A significant difference was seen in modalities for CT and sCT DRRs in longitudinal (lateral) direction, vertical direction and pitch when performing 2D lateral match. A significant difference was absorbed only in rnt when performing 2D frontal match.

Further studies are needed to obtain more detailed bone segmentation as current CT scans. Treatment planning based solely on MRI is a potential option for 2D setup verification, worth investigating further.

6.1 Future Work

This study showed MRF segmentation of MRI UTE data is useful for brain radiation therapy when performing 2D patient setup verification, but there is room for improvement of segmentation. In the future it would be interesting to use another segmentation strategy or use another MRI sequence e.g. DIXON sequence.

There is a great potential for treatment planning therapy based solely on MRI, not only for brain RT but also for the rest of the body. A T1-weighted image can be used to generate bony DRRs for 2D setup verification of e.g. prostate RT [17]. The bones appear dark in the T1-weighted images when inverting the T1 image the bony structures appear bright which will make it easier for RTTs to match with OBIs. Figure 6.1 shows T1-weighted and inverted T1-weighted image. It

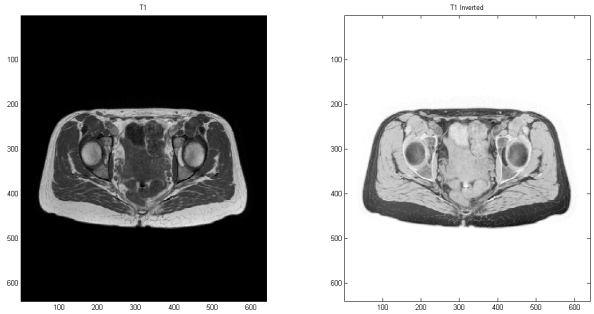


Figure 6.1: The figure illustrates cross sectional images of pelvis both for T1 and T1 inverted image, respectively.

would be interesting to make a clinical investigation of whether a significant difference in 2D patient setup verification when matching is performed with inverted T1-weighted DRRs, as compared to current CT based DRRs.

Appendix A

Appendix1

A.1 2D Manual Registration with PIPSpro

Portal Image Processing system (PIPSpro) software is an image processing system that is specially developed for portal imaging. PIPSpro measures the errors of field placement relative to anatomical landmarks set in the reference image. The registration tool in PIPSpro is used to detect and evaluate the field setup error that may occur during radiation therapy. There are three registration tools in PIPSpro that represent patient displacement with respect to the treatment field [7, 18, 6].

- Fiducial Point analysis; this registration method depends on automatic rigid-body least squares transformation where two sets of fiducial points are required.
- Template matching; is interactive matching of a reference with the features of a second image.
- Chamfer matching; is an automatic matching of either templates or fiducial points.

These registration methods will transform the treatment image into the same coordinate system as the reference image. To perform registration in PIPSpro the following steps are required [7, 6];

1. Delineation of contour fields both in reference and treatment image.

2. Defining the patientâs position using fiducial points that indicates small bone structures or template that draws over large anatomical features.

These two prerequisites must be well defined, since the relationship between contour fields and anatomical features indicates patient displacement.

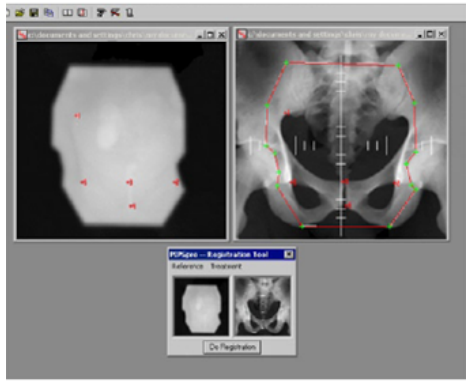


Figure A.1: The figure illustrates registration with fiducial points. This image is modified from [7].

In this study template matching registration is used, where DRR refers to reference image and OBI refers to treatment image. Template matching registration requires images with the same size. The contour field in both images must be delineated with high accuracy and anatomical features need to be drawn on reference image, which can be saved as template. During registration both contour fields are matched automatically and the template will be projected to the treatment image by using template transform control (TTC). During the registration Edit Tool will appear to scale or rotate until appropriate structure is aligned on the treatment image and the registration will be ended. The results of registration will illustrate how good the fit is accomplished. The scaling factors Mx, My and Mx/My show significant deviations of the ratio from unity that defines an out-of-plane rotation of the patient or the gantry. The field area illustrates the area of the reference, treatment and transformed treatment fields.

The transformed values Dx, Dy and Rot are achieved by aligning reference and treatment images. The values define the relative displacement and orientation of the patient features in the two images [7].

A.2 Appendix 1

A.2.1 Repeated Measures Analysis with ANOVA, Univariate and MANOVA

The section contains ANOVA tables of hypothesis which are carried out with both three-way ANOVA and MANOVA in longitudinal (lateral) direction, vertical direction, lateral direction, RNT (rotation) and pitch (rotation).

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.09219048	0.09219048	0.82	0.3773
Nurse	4	0.09123810	0.02280952	0.20	0.9345
Modality*Nurse	4	0.04876190	0.01219048	0.11	0.9784
Error	20	2.26190476	0.11309524		

Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
Frac	6	0.13780952	0.02296825	0.88	0.5104	0.4576	0.4724
Frac*Modality	6	0.07647619	0.01274603	0.49	0.8152	0.6954	0.7291
Frac*Nurse	24	0.27076190	0.01128175	0.43	0.9899	0.9461	0.9629
Frac*Modality*Nurse	24	0.33590476	0.01399603	0.54	0.9600	0.8849	0.9092
Error(Frac)	120	3.12476190	0.02603968				

Greenhouse-Geisser Epsilon	0.5125
Huynh-Feldt-Lecoutre Epsilon	0.6163

Figure A.2: The figure shows tests of hypothesis for between and within subject effects for longitudinal (lateral) direction.

The hypothesis tests for between and within subject effects at the longitudinal (lateral) are shown in figure A.2. The hypothesis for between effects are

non-significant. The univariate test shows the H-F circularity assumption is valid for this structure and the tests within subject effects are non-significant, since the adjusted H-F and G-G are above significance level. The figure A.3 The figure shows MANOVA tests of hypotheses for longitudinal (Lateral) direction. Wailk's Lambda test illustrates that the hypothesis of no fraction effect is non-significant, since the P-value is above $\alpha = 0.05$. The hypothesis for no interaction effects between fraction and modality and between fraction and nurses are non-significant. The hypothesis of no interaction effects between fraction, modality and nurse/RTT are non-significant Figure A.4 illustrates tests of hypotheses for between and within subject effects for rotation (pitch). The hypothesis test for between subject effects shows that the circulatory assumption for the structure is satisfied. The within subject effects modality, nurse/RTT and interaction between modality and nurse/RTT are non-significant. The hypothesis for within subject effects illustrates that fraction and interaction between fraction and modality are significant, while the rest are non-significant. Figure A.5 illustrates MANOVA tests of hypothesis for rotation (pitch). The hypothesis of no fraction effect is significant shown Wailk's Lambda because the P-value is below the significance level. The hypothesis for interaction no effects between fraction and modality is also significant, while the interaction no effects between fraction and nurse/RTT and between fraction, modality and nurse/RTT are non-significant.

Figure A.6 illustrates tests of hypotheses for between and within subject effects for vertical direction. The hypothesis test for between subject effects shows that modality, nurse/RTT and interaction between modality and nurse/RTT are non-significant. The hypothesis for within subject effect fraction is significant, since the P-value for the adjusted H-F is above the significance level, whereas the interactions are non-significant. The H-F and G-G circulatory assumptions for this structure are satisfied. Figure A.7 illustrates MANOVA tests of hypotheses for vertical direction. The hypothesis of no fraction effect is significant as shown in Wailk's Lambda test, since the P-value is below the significance level. The hypothesis for no interaction effects between fraction and modality and between fraction and nurses are non-significant. The hypothesis of no interaction effects between fraction, modality and nurse/RTT is also non-significant.

Figure A.8 shows tests of hypothesis for between and within subject effects for longitudinal (front) direction. The hypothesis for between subject effects modality, nurse/RTT and interaction between modality and nurse/RTT are non-significant since the P-values are above the significance level.

The hypothesis for within subject effects shows that the circulatory assumption met for this structure, since H-F ($\epsilon = 0.3115$) and G-G $\epsilon = 0.2870$ estimate of ϵ lower then 1. The two adjustments (G-G H-F) are based on a degree of freedom adjustment factor ϵ, where ϵ is satisfied when $0 < \epsilon \leq 1$. The G-G and

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac Effect
H = Type III SSCP Matrix for Frac
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.50687232	2.43	6	15	0.0762
Pillai's Trace	0.49312768	2.43	6	15	0.0762
Hotelling-Lawley Trace	0.97288343	2.43	6	15	0.0762
Roy's Greatest Root	0.97288343	2.43	6	15	0.0762

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac*Modality Effect
H = Type III SSCP Matrix for Frac*Modality
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.74813204	0.84	6	15	0.5572
Pillai's Trace	0.25186796	0.84	6	15	0.5572
Hotelling-Lawley Trace	0.33666245	0.84	6	15	0.5572
Roy's Greatest Root	0.33666245	0.84	6	15	0.5572

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Nurse Effect
H = Type III SSCP Matrix for Frac*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.30024917	0.92	24	53.539	0.5780
Pillai's Trace	0.93498817	0.92	24	72	0.5819
Hotelling-Lawley Trace	1.61631951	0.94	24	27.692	0.5541
Roy's Greatest Root	1.07554957	3.23	6	18	0.0248

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Modality*Nurse Effect
H = Type III SSCP Matrix for Frac*Modality*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.43486112	0.60	24	53.539	0.9128
Pillai's Trace	0.70127923	0.64	24	72	0.8913
Hotelling-Lawley Trace	1.00776968	0.59	24	27.692	0.9043
Roy's Greatest Root	0.62473268	1.87	6	18	0.1409

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Figure A.3: The figure shows MANOVA tests of hypotheses for longitudinal (lateral) direction.

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	4.69504762	4.69504762	1.88	0.1851
Nurse	4	25.58123810	6.39530952	2.57	0.0698
Modality*Nurse	4	5.16161905	1.29040476	0.52	0.7236
Error	20	49.84380952	2.49219048		

Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
Frac	6	23.01066667	3.83511111	4.96	0.0001	0.0017	0.0006
Frac*Modality	6	25.41295238	4.23549206	5.48	<.0001	0.0008	0.0002
Frac*Nurse	24	24.66409524	1.02767063	1.33	0.1593	0.2064	0.1843
Frac*Modality*Nurse	24	11.77038095	0.49043254	0.63	0.9011	0.8364	0.8690
Error(Frac)	120	92.69619048	0.77246825				

Greenhouse-Geisser Epsilon	0.6228
Huynh-Feldt-Lecoutre Epsilon	0.7836

Figure A.4: The figure shows tests of hypotheses for between and within subject effects for rotation (pitch).

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac Effect
H = Type III SSCP Matrix for Frac
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.36301353	4.39	6	15	0.0094
Pillai's Trace	0.63698647	4.39	6	15	0.0094
Hotelling-Lawley Trace	1.75471827	4.39	6	15	0.0094
Roy's Greatest Root	1.75471827	4.39	6	15	0.0094

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac*Modality Effect
H = Type III SSCP Matrix for Frac*Modality
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.41295865	3.55	6	15	0.0216
Pillai's Trace	0.58704135	3.55	6	15	0.0216
Hotelling-Lawley Trace	1.42154995	3.55	6	15	0.0216
Roy's Greatest Root	1.42154995	3.55	6	15	0.0216

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Nurse Effect
H = Type III SSCP Matrix for Frac*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.18710931	1.38	24	53.539	0.1651
Pillai's Trace	1.27213990	1.40	24	72	0.1392
Hotelling-Lawley Trace	2.32407736	1.36	24	27.692	0.2186
Roy's Greatest Root	1.28328462	3.85	6	18	0.0120

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Modality*Nurse Effect
H = Type III SSCP Matrix for Frac*Modality*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.42965802	0.61	24	53.539	0.9058
Pillai's Trace	0.70255011	0.64	24	72	0.8902
Hotelling-Lawley Trace	1.04262304	0.61	24	27.692	0.8895
Roy's Greatest Root	0.71764401	2.15	6	18	0.0969

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Figure A.5: The figure shows MANOVA tests of hypothesis for rotation (pitch).

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.08804762	0.08804762	1.35	0.2583
Nurse	4	0.06361905	0.01590476	0.24	0.9096
Modality*Nurse	4	0.06361905	0.01590476	0.24	0.9096
Error	20	1.30095238	0.06504762		

Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
Frac	6	0.58057143	0.09676190	7.38	<.0001	<.0001	<.0001
Frac*Modality	6	0.04361905	0.00726984	0.55	0.7654	0.6917	0.7329
Frac*Nurse	24	0.13704762	0.00571032	0.44	0.9895	0.9661	0.9814
Frac*Modality*Nurse	24	0.16638095	0.00693254	0.53	0.9636	0.9210	0.9470
Error(Frac)	120	1.57238095	0.01310317				

Greenhouse-Geisser Epsilon	0.6497
Huynh-Feldt-Lecoutre Epsilon	0.8267

Figure A.6: The figure shows tests of hypotheses for between and within subject effects for vertical direction.

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac Effect
H = Type III SSCP Matrix for Frac
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.22646548	8.54	6	15	0.0004
Pillai's Trace	0.77353452	8.54	6	15	0.0004
Hotelling-Lawley Trace	3.41568408	8.54	6	15	0.0004
Roy's Greatest Root	3.41568408	8.54	6	15	0.0004

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac*Modality Effect
H = Type III SSCP Matrix for Frac*Modality
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.85942610	0.41	6	15	0.8617
Pillai's Trace	0.14057390	0.41	6	15	0.8617
Hotelling-Lawley Trace	0.16356717	0.41	6	15	0.8617
Roy's Greatest Root	0.16356717	0.41	6	15	0.8617

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Nurse Effect
H = Type III SSCP Matrix for Frac*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.52422870	0.45	24	53.539	0.9816
Pillai's Trace	0.56675337	0.50	24	72	0.9720
Hotelling-Lawley Trace	0.74395423	0.43	24	27.692	0.9791
Roy's Greatest Root	0.44669843	1.34	6	18	0.2907

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Modality*Nurse Effect
H = Type III SSCP Matrix for Frac*Modality*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.42130513	0.63	24	53.539	0.8937
Pillai's Trace	0.71254280	0.65	24	72	0.8811
Hotelling-Lawley Trace	1.07561701	0.63	24	27.692	0.8744
Roy's Greatest Root	0.72847115	2.19	6	18	0.0928

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Figure A.7: The figure shows MANOVA tests of hypotheses for vertical direction.

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.00171429	0.00171429	0.02	0.8899
Nurse	4	0.51647619	0.12911905	1.48	0.2454
Modality*Nurse	4	0.16638095	0.04159524	0.48	0.7521
Error	20	1.74380952	0.08719048		

Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
Frac	6	0.70980952	0.11830159	1.99	0.0724	0.1574	0.1535
Frac*Modality	6	0.37495238	0.06249206	1.05	0.3962	0.3513	0.3557
Frac*Nurse	24	1.34685714	0.05611905	0.94	0.5441	0.4855	0.4893
Frac*Modality*Nurse	24	0.90361905	0.03765079	0.63	0.9025	0.7229	0.7350
Error(Frac)	120	7.13619048	0.05946825				

Greenhouse-Geisser Epsilon	0.2870
Huynh-Feldt-Lecoutre Epsilon	0.3115

Figure A.8: The figure shows tests of hypothesis for between and within subject effects for longitudinal (front) direction.

H-F adjusted P-values predict whether the effects are significant or not. For the longitudinal (front) it is seen that the within subject effects are non-significant, since these are above the significance level.

Figures A.9 illustrates repeated measures multivariate analysis for longitudinal (front) direction. Wilks' Lambda multivariate test shows that there is a little fraction effect, therefore the hypothesis of no fraction effect is rejected, since the P-value is below the significant level $\alpha = 0.05$. The hypothesis for no effects of interactions between fraction and modality and between fraction and nurse/RTT are accepted. The interactions between fraction, modality are also non-significant.

Figure A.10 illustrates tests of hypotheses for between and within subject effects for rotation (RNT). The hypothesis test for between subject effects shows that modality, nurse/RTT and interaction between modality and nurse/RTT are non-significant. The hypothesis for within subject effects illustrates that fraction is significant, while the rest are non-significant, since the adjusted H-F P-values are above significance level. The circulatory assumption for the structure is satisfied. Figure A.11 illustrates MANOVA tests of hypotheses for rotation (RNT). The Wilk's Lambda test shows that the hypothesis of no fraction effect is non-significant. This concerns interaction effects between fraction and modality and between fraction and nurses. The hypothesis of no interaction effects between fraction, modality and nurse/RTT are also non-significant.

Figure A.12 illustrates tests of hypotheses for between and within subject effects for lateral direction. The hypothesis test for between subject effects shows that modality, nurse/RTT and interaction between modality and nurse/RTT are non-significant. The hypothesis for within subject effects illustrates that fraction is significant, while interactions between fraction and modality, between fraction and nurse/RTT and between fraction, modality and nurse/RTT are non-significant. The hypothesis of no fraction effect for the lateral direction is significant shown in figure A.13, while the rest of the hypothesis tests are non-significant.

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac Effect
H = Type III SSCP Matrix for Frac
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.47119345	2.81	6	15	0.0491
Pillai's Trace	0.52880655	2.81	6	15	0.0491
Hotelling-Lawley Trace	1.12227061	2.81	6	15	0.0491
Roy's Greatest Root	1.12227061	2.81	6	15	0.0491

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac*Modality Effect
H = Type III SSCP Matrix for Frac*Modality
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.61857649	1.54	6	15	0.2316
Pillai's Trace	0.38142351	1.54	6	15	0.2316
Hotelling-Lawley Trace	0.61661494	1.54	6	15	0.2316
Roy's Greatest Root	0.61661494	1.54	6	15	0.2316

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Nurse Effect
H = Type III SSCP Matrix for Frac*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.37257281	0.73	24	53.539	0.7988
Pillai's Trace	0.77159772	0.72	24	72	0.8181
Hotelling-Lawley Trace	1.32515662	0.77	24	27.692	0.7363
Roy's Greatest Root	1.02726657	3.08	6	18	0.0296

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Modality*Nurse Effect
H = Type III SSCP Matrix for Frac*Modality*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.28697821	0.96	24	53.539	0.5290
Pillai's Trace	0.91473260	0.89	24	72	0.6142
Hotelling-Lawley Trace	1.83202860	1.07	24	27.692	0.4292
Roy's Greatest Root	1.43634437	4.31	6	18	0.0073

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Figure A.9: The figure shows MANOVA tests of hypotheses for longitudinal (front) direction.

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	5.2804286	5.2804286	0.60	0.4460
Nurse	4	11.7630476	2.9407619	0.34	0.8500
Modality*Nurse	4	1.2979048	0.3244762	0.04	0.9971
Error	20	174.7400000	8.7370000		

Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
Frac	6	26.0486667	4.3414444	3.44	0.0036	0.0500	0.0455
Frac*Modality	6	2.0252381	0.3375397	0.27	0.9512	0.7346	0.7526
Frac*Nurse	24	34.9822857	1.4575952	1.15	0.2990	0.3540	0.3521
Frac*Modality*Nurse	24	9.5380952	0.3974206	0.31	0.9991	0.9409	0.9496
Error(Frac)	120	151.6200000	1.2635000				

Greenhouse-Geisser Epsilon	0.2874
Huynh-Feldt-Lecoutre Epsilon	0.3120

Figure A.10: The figure shows tests of hypotheses for between and within subject effects for rotation (RNT).

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac Effect
H = Type III SSCP Matrix for Frac
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.65428212	1.32	6	15	0.3073
Pillai's Trace	0.34571788	1.32	6	15	0.3073
Hotelling-Lawley Trace	0.52839268	1.32	6	15	0.3073
Roy's Greatest Root	0.52839268	1.32	6	15	0.3073

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac*Modality Effect
H = Type III SSCP Matrix for Frac*Modality
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.79504776	0.64	6	15	0.6940
Pillai's Trace	0.20495224	0.64	6	15	0.6940
Hotelling-Lawley Trace	0.25778607	0.64	6	15	0.6940
Roy's Greatest Root	0.25778607	0.64	6	15	0.6940

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Nurse Effect
H = Type III SSCP Matrix for Frac*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.28527740	0.97	24	53.539	0.5226
Pillai's Trace	1.01562188	1.02	24	72	0.4532
Hotelling-Lawley Trace	1.58661386	0.93	24	27.692	0.5724
Roy's Greatest Root	0.79475116	2.38	6	18	0.0715

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Modality*Nurse Effect
H = Type III SSCP Matrix for Frac*Modality*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.37856607	0.72	24	53.539	0.8128
Pillai's Trace	0.76426132	0.71	24	72	0.8267
Hotelling-Lawley Trace	1.27885623	0.75	24	27.692	0.7644
Roy's Greatest Root	0.91518594	2.75	6	18	0.0449

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Figure A.11: The figure shows MANOVA tests of hypotheses for rotation (RNT).

Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Modality	1	0.00576190	0.00576190	0.05	0.8335
Nurse	4	0.20266667	0.05066667	0.40	0.8071
Modality*Nurse	4	0.06400000	0.01600000	0.13	0.9714
Error	20	2.54095238	0.12704762		

Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H-F-L
Frac	6	0.45523810	0.07587302	4.24	0.0007	0.0030	0.0010
Frac*Modality	6	0.10990476	0.01831746	1.02	0.4128	0.4017	0.4102
Frac*Nurse	24	0.52000000	0.02166667	1.21	0.2458	0.2742	0.2538
Frac*Modality*Nurse	24	0.35200000	0.01466667	0.82	0.7052	0.6647	0.6943
Error(Frac)	120	2.14571429	0.01788095				

Greenhouse-Geisser Epsilon	0.7008
Huynh-Feldt-Lecoutre Epsilon	0.9106

Figure A.12: The figure shows tests of hypotheses for between and within subject effects for lateral direction.

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac Effect
H = Type III SSCP Matrix for Frac
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.32272251	5.25	6	15	0.0043
Pillai's Trace	0.67727749	5.25	6	15	0.0043
Hotelling-Lawley Trace	2.09863728	5.25	6	15	0.0043
Roy's Greatest Root	2.09863728	5.25	6	15	0.0043

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no Frac*Modality Effect
H = Type III SSCP Matrix for Frac*Modality
E = Error SSCP Matrix

S=1 M=2 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.79135353	0.66	6	15	0.6834
Pillai's Trace	0.20864647	0.66	6	15	0.6834
Hotelling-Lawley Trace	0.26365773	0.66	6	15	0.6834
Roy's Greatest Root	0.26365773	0.66	6	15	0.6834

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Nurse Effect
H = Type III SSCP Matrix for Frac*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.29341832	0.94	24	53.539	0.5530
Pillai's Trace	0.93016195	0.91	24	72	0.5896
Hotelling-Lawley Trace	1.70047638	0.99	24	27.692	0.5035
Roy's Greatest Root	1.18599651	3.56	6	18	0.0168

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

MANOVA Test Criteria and F Approximations for the Hypothesis of no Frac*Modality*Nurse Effect
H = Type III SSCP Matrix for Frac*Modality*Nurse
E = Error SSCP Matrix

S=4 M=0.5 N=6.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.38147798	0.71	24	53.539	0.8194
Pillai's Trace	0.81347434	0.77	24	72	0.7648
Hotelling-Lawley Trace	1.16130580	0.68	24	27.692	0.8312
Roy's Greatest Root	0.62143094	1.86	6	18	0.1428

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Figure A.13: The figure shows MANOVA tests of hypotheses for lateral direction.

A.3 Abstract Accepted Poster Presentation at ESTRO Forum 2013

Authors: Manija Ghafory, Hans Martin Kjer, Rasmus Larsen, Jens Edmund

Objective/Purpose Radiation therapy based on MR images has proved advantageous compared to combined MRI-CT RT in terms of registration error reduction. However, lack of electron density information and MRI distortions present challenges for dose planning and generation of digitally reconstructed radiographs (DRRs) for setup verification. One option is to estimate the CT segmentation from the MR scan, a so-called substitute CT (sCT), and generate DRRs from this for bony setup verification.

In this study, we investigate whether a significant difference in 2D setup verification of a patient receiving whole brain RT could be detected when the matching was done on sCT generated DRRs as compared to normal CT based DRRs.

Material/ Method A patient receiving whole brain RT over ten fractions with 2D setup verification was investigated retrospectively. The patient data consists of a CT scan, a 1 Tesla MRI scan acquired with ultrashort echo times (UTE) and 20 anterior and lateral setup (2D) radiographs acquired at the LINAC with the On-Board Imager (OBI).

The UTE MRI was segmented into air, soft tissue and compact bone using a Markov Random Field classifier and generic HUs from ICRU report 46 to generate the sCT. The sCT was registered with the CT and the RT plan including setup fields was transferred to the sCT. The sCT DRRs were then generated in Eclipse v. 10.

Three experienced radio therapy therapists were asked to match OBIs with CT and sCT generated DRRs over the ten fractions in a random order. Matches were made with five degrees of freedom (DOF) using Offline Review with all tools available: lateral, longitudinal, vertical and two rotations rnt (anterior) and pitch (lateral). The difference in sCT- and CT-DRR based matches were treated independently for the five DOF and data from all fractions and RTTs were pooled for each DOF. A t-test per DOF was performed to determine significance ($p<0.05$) between sCT and CT based matches.

Results The t-test showed that all differences were at non-significant difference between the CT- and sCT matches for the DOFs investigated (table 1). The largest difference was seen in longitudinalLateral and lateral direction.

DOF	CT mean ± SD	sCT mean ± SD	P-value
Longitudinal$_{Lateral}$	0.1 ± 0.5	-0.7 ± 1.4	0.31
Longitudinal$_{Anterior}$	0.5 ± 0.7	0.2 ± 0.3	0.55
Vertical	0.4 ± 0.9	0.6 ± 1.0	0.73
Lateral	-1.5 ± 1.1	-1.8 ± 0.6	0.39
Pitch	0.1 ± 0.2	0.0 ± 0.0	0.61
RNT	0.2 ± 0.8	0.5 ± 0.1	0.72

Table 1: The table illustrates mean, standard deviation (SD) error in mm and P-value for longitudinal$_{lateral}$, longitudinal$_{anterior}$, vertical, lateral, pitch and RNT error for both CT-DRRs and sCT- DRRs matching.

Conclusion It was demonstrated that MRI segmented DRRs performed equally well for setup verification compared to normal CT generated DRRs showing a clinical potential for MRI only RT.

Bibliography

[1] In *http://www.varian.com/us/oncology/radiation$_o$ncology/clinac/on − board$_i$mager.html*.

[2] Sune K Buhl, Anne K Duun-Christensen, Brian H Kristensen, and Claus F Behrens. Clinical evaluation of 3d/3d mri-cbct automatching on brain tumors for online patient setup verification-a step towards mri-based treatment planning. *Acta Oncologica*, 49(7):1085–1091, 2010.

[3] Pierre Castadot, John Aldo Lee, Adriane Parraga, Xavier Geets, Benoît Macq, and Vincent Grégoire. Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors. *Radiotherapy and oncology*, 89(1):1–12, 2008.

[4] Michael J Crawley. *The R book*. Wiley, 2007.

[5] Jason A Dowling, Jonathan Lambert, Joel Parker, Olivier Salvado, Jurgen Fripp, Anne Capp, Chris Wratten, James W Denham, and Peter B Greer. An atlas-based electron density mapping method for magnetic resonance imaging (mri)-alone treatment planning and adaptive mri-based prostate radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*, 2012.

[6] L. Gilbeau, M. Octave-Prignot, T. Loncol, L. Renard, P. Scalliet, V. Grégoire, et al. Comparison of setup accuracy of three different thermoplastic masks for the treatment of brain and head and neck tumors. *Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology*, 58(2):155, 2001.

[7] Standard Imaging. *PIPSpro Manual, version 4.2*. 2008.

[8] M. Jensen and J. E. Wilhjelm. In *CT scanning, 31540 Course Note*, 2007.

[9] A. Johansson, M. Karlsson, and T. Nyholm. Ct substitute derived from mri sequences with ultrashort echo time. *Medical Physics*, 38:2708, 2011.

[10] H. M. Kjer. Mri in radiation therapy planning: The use of ultrashort echo time imaging. *Master Thesis*, 2012.

[11] K. V. Leemput and R. Larsen. *Medical Image Analysis, 02505 Course Note.* Technical University of Denmark, DTU informatics, 2011.

[12] S.Z. Li. *Markov random field modeling in image analysis.* Springer, 2009.

[13] RC Littell, PR Henry, and CB Ammerman. Statistical analysis of repeated measures data using sas procedures. *Journal of animal science*, 76(4):1216–1231, 1998.

[14] Douglas C Montgomery. *Design and analysis of experiments.* Wiley, 2008.

[15] S. Mori, T. Inaniwa, M. Kumagai, T. Kuwae, Y. Matsuzaki, T. Furukawa, T. Shirai, and K. Noda. Development of digital reconstructed radiography software at new treatment facility for carbon-ion beam scanning of national institute of radiological sciences. *Australasian Physical & Engineering Science in Medicine*, pages 1–9, 2012.

[16] S. Morris. In *Radiotherapy Physics and Equipment*, 2001.

[17] TS Nurushev, J Kim, D Hearshen, M Pantelic, M Kotys-Traughber, M Elshaikh, E Walker, I Chetty, and B Movsas. A pilot study on the feasibility of mri simulation in radiation therapy using an open-platform mr scanner. *International Journal of Radiation Oncology* Biology* Physics*, 81(2):S900, 2011.

[18] Stefania Pallotta, Marco Ceroti, and Marta Bucciolini. Accuracy of iview and pipspro registration software. *Journal of Applied Clinical Medical Physics*, 11(4), 2010.

[19] J.L. Prince and J.M. Links. *Medical imaging signals and systems.* Pearson Prentice Hall Upper Saddle River, NJ, 2006.

[20] J.A. Purdy, J.M. Michalski, J. Bradley, S. Vijayakumar, C.A. Perez, and S.H. Levitt. Three-dimensional treatment planning and conformal therapy. *Technical Basis of Radiation Therapy*, pages 179–202, 2006.

[21] J. Rahmer, P. Börnert, J. Groen, and C. Bos. Three-dimensional radial ultrashort echo-time imaging with t2 adapted sampling. *Magnetic resonance in medicine*, 55(5):1075–1082, 2006.

[22] M.D. Robson, P.D. Gatehouse, M. Bydder, and G.M. Bydder. Magnetic resonance: an introduction to ultrashort te (ute) imaging. *Journal of computer assisted tomography*, 27(6):825, 2003.

[23] JC Rosenwald, A. Nahum, and P. Mayles. *Handbook of Radiotherapy Physics: Theory and Practice*. Institute of Physics, 2002.

[24] Russ Wolfinger and Ming Chang. Comparing the sas glm and mixed procedures for repeated measures. 1998.