TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

EXPLORING NOVEL WAYS TO IMPROVE THE MRI-BASED IMAGE SEGMENTATION IN THE HEAD REGION

Jani Lindén

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 709 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2024

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

EXPLORING NOVEL WAYS TO IMPROVE THE MRI-BASED IMAGE SEGMENTATION IN THE HEAD REGION

Jani Lindén

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 709 | MATHEMATICA | TURKU 2024

University of Turku

Faculty of Science
Department of Mathematics and Statistics
Applied Mathematics
Doctoral Programme in Exact Sciences (EXACTUS)

Supervised by

Assistant Professor Riku Klén
Turku PET Centre
University of Turku
Turku University Hospital
Turku Finland

Docent Jarmo Teuho
Turku PET Centre
University of Turku
Turku University Hospital
Turku Finland

Professor Mika Teräs
Institute of Biomedicine
University of Turku
Turku Finland

Reviewed by

Professor Habib Zaidi
Geneva University Hospital
Geneva, Switzerland

Assistant Professor Jonathan Thiessen
Western University
London, Ontario, Canada

Opponent

Assistant Professor Søren Baarsgaard Hansen
Aarhus University Hospital
Aarhus, Denmark

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-9654-4 (PRINT)
ISBN 978-951-29-9655-1 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)
Painosalama, Turku, Finland 2024

To Lauren

UNIVERSITY OF TURKU
Faculty of Science
Department of Mathematics and Statistics
Applied Mathematics
JANI LINDÉN: Exploring Novel Ways to Improve the MRI-Based Image Segmentation in the Head Region
Doctoral Dissertation, 159 pp.
Doctoral Programme in Exact Sciences (EXACTUS)
April 2024

ABSTRACT

Accurate electron density information is extremely important in positron emission tomography (PET) attenuation correction (AC) and radiotherapy (RT) treatment planning (RTP), especially in the head region, as many interesting brain regions are located near the skull. Achieving good electron density information for bone is not trivial when magnetic resonance imaging (MRI) is used as a source for the anatomical structures of the head, since many MRI sequences show bone in a similar fashion as air. Various atlas-based, emission-based, and segmentation-based methods have been explored to address this problem. In this PhD project, a pipeline for MRI-based substitute CT (sCT) creation is developed and novel ways are developed to further improve the quality of bone delineation in the head region.

First, a robust sCT pipeline is developed and validated. This allows modular improvements of the various aspects of head sCT in later publications. The MRI image is segmented into different tissue classes and the final sCT image is constructed from these. The sCT images had good image quality with small non-systematic error. The time-of-flight (TOF) information improves the accuracy of PET reconstruction. The effect of TOF with different AC maps is evaluated to substantiate the need for accurate AC maps for a TOF capable system. The evaluation is performed on both subject and brain region level. While TOF information is helpful, it cannot negate the effect of the AC map quality.

The sinus region is problematic in MRI-based sCT creation, as it is easily segmented as bone. Two new methods for addressing AC in the sinus region are presented. One method tries to find the cuboid that covers the largest area of air tissue incorrectly assigned as bone and then correct the incorrect attenuation coefficient. Another method uses the sinus covering cuboid in the normalized space, from which it is converted back to each subject's individual space, after which the attenuation coefficients are calculated. Both methods improve the alignment of sCT and CT images. Finally, the possibilities of improving the quality of the bone segmentation by utilizing a random forest (RF) machine learning process is explored. The RF model is used to estimate the bone likelihood. The likelihood is then used to enhance the bone segmentation and to model the attenuation coefficient. The machine learning model improves the bone segmentation and reduces the error between sCT and CT images.

KEYWORDS: Magnetic resonance imaging, Computed tomography, attenuation correction, machine learning

TURUN YLIOPISTO
Matemaattis-luonnontieteellinen tiedekunta
Matematiikan ja tilastotieteen laitos
Sovellettu matematiikka
JANI LINDÉN: Tutkimus uusien pään alueen MRI-kuvantamiseen pohjautuvien kuvasegmentaatiomenetelmien kehittämisestä.
Väitöskirja, 159 s.
Eksaktien tieteiden tohtoriohjelma (EXACTUS)
Huhtikuu 2024

TIIVISTELMÄ

Tarkka elektronitiheystieto on hyvin tärkeää PET-kuvantamisen vaimenemiskorjauksessa sekä sädehoidon suunnittelussa erityisesti pään alueella, sillä useat kiinnostavat aivoalueet ovat kallon lähellä. Hyvän elektronitiheystiedon laskeminen luulle ei ole yksinkertaista MRI-kuvantamisen pohjalta, sillä useat MRI-sekvenssit näyttävät luut samoin kuin ilman. Useita atlas-, emissio- ja segmentointipohjaisia metodeja on tutkittu tämän ongelman ratkaisemiseksi. Tässä työssä MRI-pohjainen luotiin menetelmä MRI-pohjaisten vaihtoehto-TT-kuvien (sCT) laskemiseksi, sekä kehitetään uusia tapoja parantaa luun MRI-pohjaista erottelukykyä pään alueella.

Ensin kehitettiin ja validoitiin sCT-menetelmä. Tämä mahdollisti modulaaristen parannusten lisäämiseen sCT-menetelmään tutkimuksen myöhemmissä vaiheissa. MRI-kuva segmentoidaan eri kudosluokkiin, ja sCT-kuva lasketaan niiden pohjalta. Näin saaduissa sCT-kuvissa oli hyvä kuvanlaatu pienin epäsystemaattisin virhein. PET-kuvantamisessa fotonin lentoaikatieto (TOF) parantaa PET-rekonstruktion tarkkuutta. Tämän parantumisen määrää tutkittiin eri vaimenemiskartoilla fotonin lentoaikaa mittaavien PET-kameroiden vaimenemiskartan laatuvaatimuksien arvioimiseksi. TOF-tieto ei kokonaan pysty poistamaan vaimenemiskartan laadun vaikutusta.

Sinusten alue on ongelmallinen MRI-pohjaisessa sCT-kuvan luomisessa, sillä segmentointimenetelmät määrittävät sen usein luuksi. Kaksi uutta menetelmää esiteltiin sinusten alueen PET-kuvantamisen vaimenemiskarttojen laskentaan. Ensim-mäinen menetelmä yrittää löytää sellaisen suorakulmaisen särmiön, joka kattaisi suu-rimman mahdollisen alueen ilmaa, joka on väärin segmentoitu luuksi, ja sitten korjata tämän alueen vaimenemiskertoimen. Toinen menetelmä asettaa suorakulmaisen sär-miön normalisoituun kuva-avaruuteen, josta se käännetään takaisin kunkin henkilön omaan yksilölliseen kuva-avaruuteen, minkä jälkeen vaimenemiskertoimet määrite-tään. Molemmat menetelmät parantavat TT-kuvien ja sCT-kuvien vastaavuutta. Lo-puksi tarkasteltiin mahdollisuuksia käyttää koneoppimista ja satunnaismetsäalgo-ritmeja luun segmentoinnin parantamiseen. Satunnaismetsäalgoritmia käytetään laskemaan ennusteita kunkin kuvapisteen luutodennäköisyydelle. Luutodennäköisyyk-siä käytetään luun segmentaation parantamiseen sekä luun tiheyden arviointiin. Kone-oppimispohjainen malli parantaa luun segmentoinnin laatua, sekä vähentää virheitä TT-kuvien ja sCT-kuvien välillä.

ASIASANAT: MRI-kuvantaminen, tietokonetomografia, vaimenemiskorjaus, koneoppiminen

Table of Contents

Abbreviations

AAL	automatic anatomical labelling
AC	attenuation correction
atMRI	anatomical MRI
AUC	area under curve
CG	cuboid goodness
COM	centre-of-mass
CSF	cerebrospinal fluid
CT	computed tomography
CTAC	CT-based attenuation correction
DRF	distributed random forest
FDG	fluorodeoxyglucose
FEG	frequency encode gradient
FID	free induction decay
FPB	filtered back projection
FWHM	full width at half maximum
GM	grey matter
HU	Hounsfield unit
ICBM	International Consortium for Brain Mapping
LOR	line of response
MAE	mean average error
ML	machine learning
ML-EM	maximum likelihood expectation maximization
MNI	Montreal Neurological Institute
MRI	magnetic resonance imaging
MR-RT	magnetic resonance radiotherapy
NAC-PET	non-attenuation corrected PET
OSEM	ordered subsets expectation maximization
PEG	phase encode gradient
PET	positron emission tomography
RAMLA	row-action maximum likelihood algorithm
RF	random forest

ROC	receiver operating characteristics
RT	radiotherapy
RTP	RT treatment plans
sCT	substitute CTs
SPECT	single photon emission computed tomography
SSD	square sum of distances
SSG	slice select gradient
TPB-AC	tissue probability based AC
TOF	time-of-flight
TPM	tissue probability map
TXAC	transmission-based attenuation correction
UTE	ultra-short echo time
VOI	volume of interest
WM	white matter
XRT	extremely randomized trees
ZTE	zero echo time

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I J. Teuho, J. Linden, J. Johansson, J. Tuisku, T. Tuokkola, and M. Teräs. Tissue Probability-Based Attenuation Correction for Brain PET/MRI by Using SPM8. *IEEE TRANSACTIONS ON NUCLEAR SCIENCE*, 2016; Vol 63: 2452–2463. © 2016 IEEE.

II Lindén J, Teuho J, Klén R, Teräs M. Are Quantitative Errors Reduced with Time-of-Flight Reconstruction When Using Imperfect MR-Based Attenuation Maps for 18F-FDG PET/MRI Neuroimaging, Applied Sciences. 2022; 12(9):4605. https://doi.org/10.3390/app12094605

III Lindén, J., Teuho, J., Teräs, M., Klén R. Evaluation of three methods for delineation and attenuation estimation of the sinus region in MR-based attenuation correction for brain PET-MRI imaging. BMC Med Imaging 22, 48 (2022). https://doi.org/10.1186/s12880-022-00770-0

IV Lindén J, Ranta I, Ramanayake Mudiyanselage K, Teräs M, Keyriläinen J, Teuho J, Klén R. Creation of bone tissue probability maps which correlate with bone density from the head and neck region by machine learning with sparse datasets. *Submitted for peer review in Medical & Biological Engineering & Computing*.

The original publications have been reproduced with the permission of the copyright holders.

1 Introduction

Often it is medically interesting or necessary to see what is under the skin of the patient. For a long time, the only way to see the inner anatomy of a patient was to perform an invasive explorative surgery. This changed in 1895 when Wilhelm Röntgen discovered X-rays (Röntgen, 1896; Rosenow, 1995; Nüsslin, 2020). Since then, the imaging field has expanded to encompass a large variety of different modalities. In particular, X-rays were discussed to have therapeutic effects and were used for radiotherapy (RT) as early as 1896 (Bowles, 1896; Grubbe, 1906; Grubbé, 1933; Lederman, 1981; Holsti, 1995).

1.1 CT and MRI

While positron emission tomography (PET) measures the physiology of the subject, computed tomography (CT) and magnetic resonance imaging (MRI) are imaging modalities that are more focused on the anatomy of the subject. In CT, an X-ray tube and detectors are mounted on a gantry where they can rotate around the z-axis of the CT scanner (Bushberg *et al.*, 2011, ch 10). This allows projection images to be taken from any angle around the patient. Modern CT scanners have the detector array arranged as an arc, so the fan of X-rays produced by the X-ray tube all hit the detector approximately orthogonally. When the patient has been imaged the photon intensities around the body in all table positions are ordered into projection distributions and then reconstructed into a tomographic image using computer algorithms.

Since X-rays utilize ionizing radiation, each scan statistically increases the chances of a patient developing cancer. The usage of MRI avoids exposure to ionizing radiation and allows more flexibility in imaging various tissues. An MRI scanner has a large gantry which is magnetized with strong a magnetic field (1.5T and above) (Bushberg *et al.*, 2011, ch 12).

The magnetic field will align the atoms with non-zero spin, such as hydrogen atoms, in the patient's body parallel to the magnetic field. Hydrogen atoms, or just protons, have the strongest magnetic moment, and are abundant in water and fat, which makes them the principal source of MRI signals. While most protons don't

follow the direction of the magnetic field, enough do so that the difference in net magnetization is measurable. The protons do not align perfectly with the magnetic field but do have their rotational axis wobbling around the direction of the magnetic field; and the frequency of this precession is called Larmor frequency. The protons' precession is out of phase with other protons but applying a Larmor frequency radio signal transversely to the direction of the magnetic field causes the protons' rotational axis to align with each other. This creates transverse magnetization and depending on the angle of the radiofrequency pulse and flip angle, it reduces the longitudinal magnetization. When the protons' precession is in sync, the rotation of the protons will induce a radiofrequency signal (free induction decay, FID) and as the dephasing occurs the FID is reduced. The FID signal intensity can be estimated from the equation (Serai, 2022)

$$S(T_e) = S_0 e^{-\frac{T_e}{T2}} \tag{1}$$

where $S(T_e)$ is the signal intensity, measured at echo time T_e, S_0 is the initial signal intensity and $T2$ is a tissue specific value.

The time for the FID to reach $1/\mathrm{e} \approx 37\%$ of the original amplitude is called the T2 relaxation time and it varies between different tissue types. However, due to extrinsic variations in the magnetic field, or different susceptibility agents in the tissues, more rapid dephasing will occur. This is referred to as T2*.

At the same time with the decline of transversal magnetization, the longitudinal magnetization is recovered when the protons spins are aligned again with the main magnetic field. The recovery of the magnetization can be estimated using equation (Serai, 2022)

$$M(t) = M_0 \left(1 - e^{-\frac{t}{T1}}\right) \tag{2}$$

where $M(t)$ is the magnetization, measured at time t, M_0 is the original magnetization, and $T1$ is a tissue specific value.

This is called the T1 relaxation, and the time for the recovery of $1\text{- } 1/\mathrm{e} \approx 63\%$ of the longitudinal magnetization is called the T1 time. Since different tissues have different T1 times, it is possible to produce different T1-weighted MRI images by applying the radiofrequency signals repeatedly. A radiofrequency signal that is antiparallel to the main magnetic field can be applied, which effectively reverses the longitudinal magnetization. Different tissues then start returning to the equilibrium at different paces, and if a transversal radiofrequency pulse is applied when the net longitudinal magnetization of some tissue is at zero that tissue does not produce any FID, creating MRI images with no signal from one tissue class.

The whole area of interest is not excited at once, as that would not yield any anatomical resolution. Instead, several gradients are applied to the patient to localize the signal spatially. First is the slice select gradient (SSG), and it changes the strength of the magnetic field along the cranial – caudal axis. Because the Larmor frequency is dependent on the strength of the field, SSG makes each radiofrequency pulse to affect only one slice of the patient. Steeper the SSG, thinner the resulting slices are. Next a frequency encode gradient (FEG) is applied perpendicularly to the SSG. The alignment of FEG sets the x-axis of the image. The FEG is applied around the time of echo signal, and it changes the frequency of the FID along the x-axis. Now the receiving antenna gets signals with slightly different frequencies, and the amplitude of each position along x-axis can be determined with signal processing. Finally, a phase encode gradient (PEG) is applied perpendicular to both SSG and FEG. PEG determines the position of the protons on y-axis. PEG is applied after the SSG but before the FEG. It will change the phase of the FID on different positions along the y-axis.

The signals measured are saved on a matrix called k-space matrix. The middle of the k-space has the lowest frequency variation and produces the main shapes of the final image. Further from the centre are the higher frequency variations and they yield the finer details of the final image. The k-space is symmetric across the midpoint, so once half of the rows plus one row of the desired data is gathered, the other half can be synthesized (Bushberg *et al.*, 2011, ch 13). Depending on the scan, the full k-space might not be filled, leaving part of it empty. Once k-space is filled sufficiently, an inverse Fourier-transform can be performed to take the frequency space image into the spatial space for viewing.

1.2 PET and SPECT Imaging

X-ray and CT imaging are mostly anatomical, while some functional applications exist. To get fully functional images, a different type of modality is needed. Nuclear medicine comprises two major imaging modalities known as single photon emission computed tomography (SPECT) and PET. Both of these use radioactive decay of certain elements to create a picture of the metabolism and functionality of a patient (Bushberg *et al.*, 2011, ch. 18 and ch 19). In PET, atoms of radioactive isotopes are tagged into molecules that will, once injected into the patient, behave like similar molecules already existing within the body, or have otherwise useful metabolic or accumulative characteristics. The radionuclides used in PET imaging decay with positron emission. A positron will travel a short distance in the tissue, and once it has shed most of its kinetic energy it will interact with an electron annihilating both and generating two 511 keV gamma photons traveling to opposite directions (Figure 1).

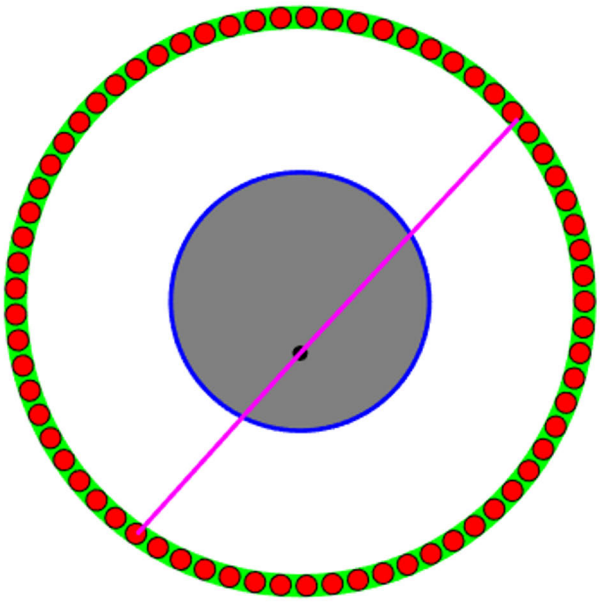

Figure 1. Diagram of PET imaging. A radiotracer molecule beta decays, and the resulting positron meets an electron (black dot). The annihilation gamma rays (purple) are detected by the photomultipliers and scintillators (red) on the scanner gantry (green). The line between the green detectors is called line-of-response.

The gamma photons are then detected by the PET gantry almost simultaneously. Since the photons are traveling to opposite directions, the decay of the positron must have happened between the two detectors that detected the gamma photons. This line between the two detectors is called the line of response (LOR). When two of these signals are detected within a pre-set time window, it is considered a coincident event and logged. The time window depends on the specifications of the computer system and scanner. The shorter the time window, the less likely it is to log a random incident as a true incident. A random incident is a detection of a photon pair that is not a result of one annihilation along a LOR, but two annihilations from both of which only one photon is detected. After each detection, the detector needs to wait before it can detect another photon. This phenomenon sets the upper limit on the amount of dose that can be injected to the patient if radiological safety was not an issue. Alongside randomly detected and true incidents, the scanner can also detect scattered incidents. These happen when one or both gamma rays scatter, or change direction, when traveling through the tissues. This results in scattered incidents to be logged into an incorrect LOR. Therefore, they need to be corrected for using statistical methods.

SPECT is a functional modality like PET, but instead of positron emitting isotopes, SPECT uses isotopes that emit gamma photons, such as Technetium-99m,

Iodine-123, or Thallium-201. These can also be tagged onto molecules to better mimic various metabolic processes. Whereas PET relies on the detection of the annihilation photons, which always have the energy of 511 keV, SPECT can use a combination of gamma emitting tracers simultaneously, provided that the photon energies are distinct enough that the system can separate them.

SPECT system has one or multiple camera heads that rotate around the subject acquiring projection images over time. As SPECT uses single photons to record incidents, the camera heads include collimators that narrow the arc from which the incident photon can reach the camera, thus improving resolution.

When enough photons, or photon pairs for PET, are detected an image of the underlying radioactivity distribution can be reconstructed. The incidents are organized into projection images, or sinograms, where each pixel of the image corresponds to a certain LOR. Before reconstruction, an attenuation correction needs to be applied. Because the gamma photons attenuate inside the patient passing through different tissues, the signal is weakened depending on the tissues through which the photons pass. Accounting for this is of the utmost importance, so that the exact activity in the source tissue can be quantified.

The traditional way to estimate the attenuation in PET is to rotate a rod source of gamma rays around the patient and measure the attenuation with this transmission data. During early times with 2D PET imaging, the scattered photons were not a very large fraction of all detected photons, and this allowed for straightforward quantification of activity (Bailey and Willowson, 2013). These days it's more common to estimate the attenuation with a CT or MRI scanner. This complicates quantification and causes PET imaging to face similar quantification issues SPECT has been facing since its inception. Often the PET and SPECT scanners are coupled physically with the CT scanner making it easier to co-register the emission data with the attenuation correction (AC) data. After the projection data has been corrected for attenuation, the image of the underlying radioactive distribution can be reconstructed using various algorithms.

1.2.1 Development of PET scanners

PET scanners have undergone significant developments since they were first introduced in the 1970s. In the beginning, PET was primarily used for research purposes, and the first human PET tomography was performed in 1974 (Nutt, 2002). Since the 1980s, PET imaging started to be used more widely for clinical purposes, and it was used to study a variety of medical conditions, including cancer, heart disease, and brain disorders (Nutt, 2002; Alauddin, 2012; Ahuja *et al.*, 2020; Davis *et al.*, 2020).

In the 1990s, PET imaging technology continued to improve, and the use of PET scans for diagnostic and therapeutic purposes became more widespread. Advances in computer technology also made it possible to more easily process and interpret the large amounts of data produced by PET scans, which further increased their usefulness in clinical practices. In the last couple decades, the hybrid modality scanners started to become more prevalent. PET scanners were combined with both CT (Townsend *et al.*, 1999) and MRI (Shao *et al.*, 1997). These hybrid scanners made it easier to combine anatomical information to the functional image produced by the PET scan.

PET-CT devices can be particularly useful in cancer diagnosis and treatment, as it allows for the localization of abnormalities detected on the PET scan within specific structures within the body. In the meanwhile, PET-MRI devices are more effective, for example, in brain imaging, where the combination of PET and MRI can provide more detailed and complementary information about brain function and structure. In recent years, PET imaging has been used to guide the delivery of radiation therapy, helping doctors to target cancerous tumours while minimizing the risk of side effects more accurately to healthy tissues.

The resolution of PET scanners has improved significantly over the years, as there has been advances in detector technology and image reconstruction algorithms. Early scanners had relatively low spatial resolution, which limited their ability to accurately depict small structures in the body, but the newer PET scanners' improved resolution allows for more detailed images of the body's functions and metabolism. One field of improvement is time-of-flight (TOF) technology. Being able to locate the place of annihilation more accurately than any point between the two detectors that detected the incident would be beneficial for signal to noise ratio of the PET imaging (Wong *et al.*, 1983; Vandenberghe and Karp, 2006; Conti, 2009). The first TOF PET scanners were developed in the 1980s (Surti and Karp, 2016). Since then, the time resolution of the TOF capable PET scanners have improved, and there are now scanners with time resolution near 200ps on the market (Van Sluis *et al.*, 2019).

1.2.2 Radioligands

PET uses radioligands, also known as radiotracers or tracers, to produce images of the body's biochemical and physiological processes. Radioligands are molecules that are labelled with a radionuclide, a type of radioisotope that decays by emitting positrons. The radioisotope can be one of two kinds of isotopes. One being an isotope of the atom that is normally on the molecule, such as, Oxygen-15 in O_2 or H_2O (Clark *et al.*, 1987). Otherwise, it would be an isotope of another atom that behaves similarly enough to be applicable, such as Fluoride-18 in fluorodeoxyglucose (FDG) (Alauddin, 2012).

Radioligands are used in PET to study a wide range of biological processes, including brain function, cancer metabolism, and heart function. They are also used to diagnose and monitor a variety of medical conditions, such as Alzheimer's disease, epilepsy, inflammation, and Parkinson's disease (Love and Palestro, 2004; Brooks, 2010; Jack *et al.*, 2010; Kumar and Chugani, 2013).

There are many different radioligands used in PET, each with specific properties that make it suitable for different applications. For example, some radioligands are able to cross the blood-brain barrier and accumulate in the brain, making them useful for imaging brain function. Other radioligands are designed to bind to specific receptors or enzymes, enabling them to target specific tissues or organs.

The tracer compounds used in PET scans have undergone significant development over time. Due to the initial stages of radioligand development, the early PET scanners had to use a relatively small selection of tracers, such as FDG (Ido *et al.*, 1978), or some other Fluoride-18 and Carbon-11 labelled ligands like [^{11}C] raclopride (Farde *et al.*, 1985), [^{11}C] methionine (Comar *et al.*, 1976), [^{18}F]Fluorodopa (Garnett, Firnau and Nahmias, 1983) and, [^{11}C]Carfentanil (Frost *et al.*, 1985). As radiochemistry has progressed, the more recent PET scanners have a wider range of tracer compounds to choose from, facilitating for more detailed and specific imaging.

FDG is a sugar analogue that is commonly used in PET imaging to visualize glucose metabolism in the body. It is often used to detect and monitor cancer, as many types of cancer cells have high rates of glucose metabolism. Once the FDG enters the cell, it cannot be processed further until the radioactive decay, as it is missing a hydroxyl group. After the positron emission the fluoride atom becomes an Oxygen-18 ion, which quickly picks up a proton from its aqueous environment turning the molecule into a normal glucose that can then be metabolized normally.

Myocardial perfusion PET uses a few radioligands to study the heart function. The half-lives of these are short, so their production must be arranged close to the imaging site, and scan times are short. Some ligands, such as radioactive water (using Oxygen-15) and ammonia (using Nitrogen-13) can be synthesized with the help of cyclotrons, while Rubidium-82 is produced by a Strontium-82/Rubidium-82 generator.

For neurological studies there are several groups of radioligands available. Dopamine transporter radioligands, such as [^{18}F]CFT (Haaparanta *et al.*, 1996) and [^{18}F]FP-CIT (Goodman *et al.*, 1997), are used to visualize the dopamine system in the brain, which is involved in reward and motivation. Serotonin transporter radioligands, like [^{11}C]DASB (Wilson *et al.*, 2000), are used to visualize the serotonin system in the brain, which is involved in mood and emotional regulation. GABA-A receptor radioligands, such as [^{18}F]flumazenil (Ryzhikov *et al.*, 2005) and [^{11}C]flumazenil (Persson *et al.*, 1985), are used to visualize the GABA-A receptors

in the brain, which are involved in anxiety and stress. Amyloid-beta (Aβ) radioligands, such as [^{18}F]flutemetamol (Koole *et al.*, 2009) and [^{11}C]PIB, are used to visualize the accumulation of amyloid-beta in the brain, which is a hallmark of Alzheimer's disease.

Development of new tracers occurring constantly. Fortunately, the newer generation tracers often overcome certain problems of earlier tracers. These problems include protein binding, uptake in unintended organs, or too quick excretion through kidneys.

1.2.3 PET reconstruction algorithms

PET reconstruction can be done analytically or iteratively. In analytical algorithms, the image is back projected from the measured projections, but due to inherent noise in the physical measurements, using simple back projection creates a very blurry image. Instead, the projection data is transformed into Fourier space, where a filter is applied damping the high frequency parts of the signal. The filtered signal is transformed back to the projection space and then back projected to form the image. This algorithm is called filtered back projection (FPB) and it's fast but has some artefacts since the projection angles are not infinitely close to each other. Iterative algorithms start from a guess for the image; a predetermined shape, a uniform field, or even result of the FPB. The image is then projected into the same projection angles as the measurement data is in, and the differences between projections are compared. The initial image is then adjusted, and a new projection data is calculated. This process is continued until a set threshold on the difference between image projection and measurement data has been reached, or a set number of iterations are completed. Iteratively calculated images are typically smoother, and the iterative algorithms can consider physical characteristics of the measurement in the projection phase. However, they take longer than FBP to calculate, but if quantitative accuracy is required, then fully converged iterative reconstruction, or analytical reconstruction, is the more appropriate choice (Reader and Zaidi, 2007).

The algorithms in this study for PET reconstruction are ordered subsets expectation maximization (OSEM) and row-action maximum likelihood algorithm (RAMLA). Both algorithms are iterative and based on the maximum likelihood expectation maximization (ML-EM) algorithm (Shepp and Vardi, 1982):

$$\mathrm{X}_v^{(k)} = \frac{\mathrm{X}_v^{(k-1)}}{\sum_{i=1}^{n} a_{iv}} \sum_{i=1}^{n} a_{iv} \frac{b_i}{\langle A_i, X^{(k-1)} \rangle}, \tag{3}$$

where X is the image vector, k is the iteration round, $v \in \{1, \ldots, m\}$ is the index for all elements in X, a is an element in the projection matrix A, $i \in \{1, \ldots, n\}$ is the index for a LOR, and a column in the projection matrix A with n LORs and finally

b_i is an element in the measurement vector b. For now, the elements in the image vector are considered to be voxels, or non-overlapping cubes filling the image space.

Since the ML-EM algorithm always uses voxel values from the previous iteration to calculate the voxel values for the next iteration. This means the convergence of the image vector towards the final image takes many iterations. The rate of convergence could be improved by updating the image vector several times within one iteration.

Let now k be the index for full iteration round, $q \in \{1, \dots, p\}$ be the index for sub cycle, and S_q be the subset of LORs that belongs to sub cycle q. The sub cycles and iterations can be connected by the notations $X^{(k,0)} = X^{k-1}$ and $X^{(k,p)} = X^k$. The algorithm completes one full iteration when all the sub cycles have been processed. A sum term

$$c_{qv} = \sum_{l \in S_q} a_{lv} \tag{4}$$

corresponds to the projection matrix column sum

$$\sum_{i=1}^{n} a_{iv} \tag{5}$$

used in the ML-EM algorithm. Now the ML-EM algorithm can be written in OSEM form (Hudson and Larkin, 1994)

$$X_j^{(k,q)} = \frac{X^{(k,q-1)} d_{qj}}{c_{qj}} \tag{6}$$

where the term

$$d_{qv} = \sum_{l \in S_q} \frac{a_{lv} b_l}{\langle A_l, X_v^{(k,q-1)} \rangle} \tag{7}$$

for indices $v = 1, \dots, m$ and $q = 1, \dots, p$. The vector A_l is a column in the projection matrix and the term b_l is the measured value for LOR l. All the voxel values in the image vector are updated with each sub cycle q. The term d_{qv} remains the same for all elements within a sub cycle.

The OSEM algorithm can be taken further by setting the number of sub cycles to be equal to the number of LORs. This, with some other adjustments, will become the RAMLA formula. Let A be the projection matrix, and a_{ij} its elements. Let $X^0 > 0$ be an initial guess for an image vector and relaxation parameter $\lambda_k > 0$ so that $\lambda_k a_{ij} \leq 1$ for all $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. The relaxation parameter λ_k should satisfy the conditions.

$$\begin{cases} \lim_{k\to\infty} \lambda_k = 0, \\ \sum_{k=1}^{\infty} \lambda_k = \infty. \end{cases} \tag{8}$$

In practice the calculations are stopped before infinity. Thus, a clear decrease in the value of λ_k is sufficient to reconstruct an image of a good quality.

Like in OSEM case, let k denote the full iteration round and $i \in \{1, \dots n\}$ index for the sub cycle, or LOR being processed. Again, the sub cycles and iterations are connected by the notations $X^{(k,0)} = X^{k-1}$ and $X^{(k,n)} = X^k$. The algorithm will complete one full iteration when all the LORs have been processed. The RAMLA formula is derived from the OSEM formula by first setting the relaxation parameter $\lambda_k^i = 1/c_{ij}$ and increasing the sub cycle count to match the number of LORs. This will lead to formula

$$X_j^{(k,i)} = X_j^{(k,i-1)} \lambda_k^i d_{ij} \tag{9}$$

where

$$d_{ij} = \frac{a_{ij} b_i}{\langle A_i, x^{(k,i-1)} \rangle}. \tag{10}$$

Now $1/c_{ij} = 1/a_{ij} = \lambda_k^i$ or equally $\left(1 - a_{ij}\lambda_k^i\right) = 0$ so the previous formula can be written as

$$X_j^{(k,i)} = X_j^{(k,i-1)} \left(1 - a_{ij}\lambda_k^i\right) + X_j^{(k,i-1)} \lambda_k^i d_{ij} \tag{11}$$

where d_{ij} is same as above.

The first term, which still gets the value zero, is needed later so that all image vector values do not approach zero when the relaxation parameter λ approaches zero, when the number of iterations increases. Let the relaxation parameter λ be fixed as a constant for all sub cycles $i \in \{1, \dots n\}$ and match the conditions set earlier. This results in RAMLA formula

$$X_j^{(k,i)} = X_j^{(k,i-1)} \left(1 - \lambda_k a_{ij}\right) + X_j^{(k,i-1)} \lambda_k a_{ij} d_i \tag{12}$$

where similarly as in Eq. 8

$$d_i = \frac{b_i}{\langle A_i, X^{(k,i-1)} \rangle} \tag{13}$$

and λ_k remains a constant through all sub cycles $i \in \{1, \dots n\}$. As the image vector is updated after each LOR, the order of LOR processing makes a difference to the end result if the image is not allowed to converge to infinity. The term d_i depends only on the LOR, and not the voxel so term d_i can be used for all voxels within the

sub cycle. This speeds up the calculations when compared to the earlier OSEM formula.

In the original publication (Browne and De Pierru, 1996), the RAMLA formula was written in an equivalent form

$$X_j^{(k,i)} = X_j^{(k,i-1)} + \lambda_k X_j^{(k,i-1)} a_{ij} \left(\frac{b_i}{\langle A_i, X^{(k,i-1)} \rangle} - 1 \right). \quad (14)$$

The part within the parenthesis remains constant for all the voxels in the image. And the formula shows the purpose of the RAMLA approach. The voxel value is the first term in the formula. Then the relative difference between the current estimate and the measured LOR value is calculated and scaled to the voxel in question using the factor a_{ij}. Then this is multiplied with the voxel value to get the absolute difference. Now importantly, the relaxation parameter λ_k limits how much of the absolute difference between the actual LOR measurement and the current estimate is passed on to the voxels along the LOR. This allows for smoother development of the image when the reconstruction proceeds. Eventually when the relaxation parameter λ_k approaches zero the convergence is achieved.

Traditionally the image space is split into voxels that are cubes filling the space without any overlap. Any point in the image space belongs to exactly one voxel, so the center point of the voxel can be considered to represent all the other points within the voxel and connect to the value in the image vector. This also makes it straightforward to convert voxel image space to images or pictures for further calculation or expert viewing. However, the voxels are not necessarily the ideal base function for dividing the image space. The non-overlapping nature of the voxels can cause them to create large differences between neighboring voxels, or noise in the image. Additionally, as there can be no overlap, a cartesian grid is most feasible with cube voxels.

One alternative for voxels is overlapping spherical volume elements, or blobs. Blobs have been studied for a long time (Lewitt, 1990, 1992; Matej and Lewitt, 1995, 1996), and offer some interesting mathematical benefits at the cost of increased complexity of reconstruction. They are arranged in grids as well, but since there is overlap, they can be arranged differently, such as body-centered cubes or face-centered cubes. One PET scanner used in this study uses blobs instead of voxels in the reconstruction algorithms. Unlike voxels that have base a function value of 0 or 1, depending on whether the point in the image space is within the voxel or not, the blobs have a continuous value range between 0 and 1. The value of one is reached only in the middle of the blob, and value of 0 is reached at the edge of the blob and beyond. The blobs are defined as

$$b_{m,a,\alpha}(r) = \begin{cases} \frac{1}{I_m(\alpha)} \left[\sqrt{1-(r/a)^2}\right]^m I_m\left[\sqrt{1-(r/a)^2}\right], 0 \le r \le a \\ 0, r > a \end{cases} \quad (15)$$

where I_m is the first type modified Bessel function of degree m, $a > 1$ is the radius of the blob, α is a parameter controlling the shape of the blob, and r is the distance from the center of the blob. If $m = 0$ then the blob is not continuous when $r = a$. If $m > 0$, then the blob has $m - 1$ continuous differentials when $r = a$. If the shape parameter α is small, then the value of the blob takes longer to decrease while large values of parameter α result in blobs that decrease faster in value. Typically, the reasonable values of α and a depend on the characteristics of the image and data, but generally the radius a should be picked in a way that the full width at half maximum (FWHM) of the blobs does not exceed that of the PET scan. Breaking this criterion can lead to overestimation of contrast in high contrast areas and artefacts in the resulting image.

There are combinations of parameters α and a that result in better quality images than some other combinations. The exact optimal parameters depend on the grid structure (Li, 2018). Using blobs as a base function for reconstruction allows for more dynamic selection of parameters affecting image quality. Different values for parameters α and a can be chosen based on the individual situation and imaging conditions.

1.2.4 Time of Flight

Originally, any two photons arriving at two detectors close enough to each other were marked as an incident on that specific LOR. It was not possible to estimate the point of annihilation along the LOR. When the timing resolution of the PET detectors increased, this changed. As the annihilation photons in PET travel at the speed of light, there is a time between the photons reaching the different detectors on either end of the LOR where the annihilation happened, if the annihilation did not occur precisely on the midpoint of the LOR. If the time difference between the photons arriving at the detectors can be measured, the location of the annihilation can be calculated (Bailey *et al.*, 2005, ch 1).

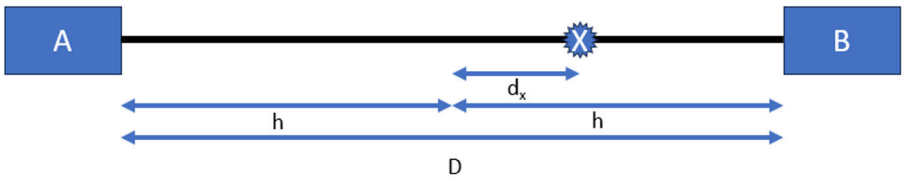

Figure 2. Illustration of TOF measurement. X marks the point along the LOR where the annihilation happens, and the photons originate from. Photons are detected at detectors A and B.

The derivation of the calculation can be done based on Figure 2. Point X marks the annihilation. If point X was at the halfway of the distance D, then both photons would travel half the distance h until they reach the detector. Now that point X is not at the halfway point, one photon travels the distance $h - d_x$ and the other travels the distance $h + d_x$ before reaching the detector. The extra distance $2d_x$ causes there to be a time difference t between the arrival times at the detectors A and B. The distance from the midpoint can then be calculated as $d_x = ct/2$, where c is the speed of light. The accuracy of distance d_x depends on the timing resolution of the PET system. Timing resolution at almost 200ps has been reported (Van Sluis *et al.*, 2019) and that corresponds to an inaccuracy of around 6cm.

Storing the TOF data is more complex than it is for non-TOF measurements due to the annihilation locations being variable. An intuitive approach is to store the measurements in list mode, but this can lead to slow processing times (Vandenberghe *et al.*, 2016). Another approach is to use multiple sinograms, each for a different segment of the LORs between the start and end detectors.

While performing image reconstruction, assigning the point of annihilation more accurately along the LOR allows for updating only the values of the voxels that are located on the segment of the LOR where the annihilation could have happened. This allows the reconstruction to converge faster (Mehranian and Zaidi, 2015b). Furthermore, as the cross dependencies between the voxels are reduced and the additional constraints brought by the more accurate location estimate of the annihilations, the reconstruction is less prone to errors in AC, and scatter correction (Conti, 2011; Conti and Bendriem, 2019). Additional benefits of TOF reconstruction are better contrast and signal-to-noise ratio (Conti, 2009), and reduced errors due to faulty PET detectors or various artefacts and implants (Davison *et al.*, 2015; Delso *et al.*, 2017; Svirydenka *et al.*, 2017; ter Voert *et al.*, 2017).

To use TOF effectively, some adjustments need to be applied to the measurements (Vandenberghe *et al.*, 2016). The first problem to correct is a timing offset that causes different detectors to detect the photons at slightly different times because of the detector and electronic setup of the system. Secondly, the probability distribution of the annihilation point, or the TOF kernel, needs to be estimated. Incorrect kernel can cause artefacts in the reconstructed image.

TOF offers possibilities that were not as viable for scanners without the TOF functionality (Vandenberghe *et al.*, 2016). As the TOF measurements are more accurate, the target regions of interest can be made much smaller than with non-TOF scanners. Also, radioisotopes where only a small share of decays lead to positron emission can become more viable to use, as TOF reconstruction is more sensitive than non-TOF reconstruction. Overall, if the TOF reconstruction is well calibrated, it can be expected to provide better contrast and signal-to-noise ratio than its non-TOF counterpart.

1.3 Radiotherapy

Radiotherapy (RT) works the opposite way to PET from outside in. Typically, a tumour is found within the subject, and it is decided that radiating the tumour, potentially alongside other treatment options, is a prudent course of action. A suitable dose is calculated, and the subject is subjected to radiation that penetrates through the tissue and damages the tumour. Again, as the radiation travels through the tissue, it is important to account for the different tissues it passes through, so that the tumour gets the correct amount of radiation. Too little might not achieve the treatment goals, and too much will inevitably damage more of the surrounding tissue.

The use of radiation to treat disease dates to the late 19th century, when it was discovered that certain types of radiation could be used to kill cancer cells. X-rays were first used in radiotherapy, soon after their invention, in 1896 to treat carcinoma of the breast and lupus vulgaris (Lederman, 1981). Other early applications of radiotherapy were in the treatment of cancers on skin and body cavities such as the cervix with either X-rays or placing radium tubes in contact with the tumors directly. These earlier forms of radiotherapy had limited effectiveness and were associated with significant side effects.

After early experiments, radiotherapy developed into more of a scientific technique along with the work of Regaud, Coutard and Hautant in 1920s and 30s (Holsti, 1995). In the mid-20th century, advances in physics and technology led to the development of more sophisticated forms of radiotherapy which are typically divided into three types. First of which is external beam radiotherapy, which uses linear accelerators to deliver high-energy (4 to 25 MeV) radiation to the cancerous area from outside the body (Klein *et al.*, 2009). The second is brachytherapy, which involves the insertion of small radioactive objects into the tumor, for example in the cervix (Nag *et al.*, 2000). Lastly, systemic radioisotope therapy where the radioisotopes are given in therapeutic amounts by oral ingestion or infusion, for example for prostate cancer with bone metastasis (Sartor *et al.*, 2017).

Originally, treatment planning was done using two dimensional radiation therapy approach with an X-ray simulator (Purdy, 2008). Once CT became more widely available, it was applied also to RT, to help transition towards 3D planning. CT allowed for a much better understanding of the volume of the target tumour and, thusly, helped to increase the quality of treatment simulator software (Goitein and Abrams, 1983). CT also enabled the use of non-coplanar beam angulations. Although bringing great improvements in the treatment quality, CT is not always able to alone identify the gross tumour volume accurately so other modalities, like MRI, or nuclear medicine would be needed to assist in tumour size evaluation (Purdy, 2001).

An important aspect of improved treatment planning is magnetic resonance radiotherapy (MR-RT). In MR-RT, MRI is combined with external beam

radiotherapy whereby a patient undergoes an MRI scan to create detailed images of the tumour and surrounding tissue. These images are then used to guide the delivery of external beam radiotherapy to the tumour. MR-RT can be used alongside other imaging modalities like CT, as MR alone might not be comprehensive enough to accurately plan the RT. MR-RT may also be more effective in treating certain types of tumours, such as those that are difficult to visualize on other types of imaging. While MR-RT is a relatively new technology, there has been extensive research on the usage of MRI-based treatment planning (Metcalfe *et al.*, 2013; Owrangi, Greer and Glide-Hurst, 2018).

There are also other applications of MRI used in conjunction with radiotherapy. In MRI-guided radiotherapy, MRI is used to guide the delivery of radiotherapy in real-time (Pollard *et al.*, 2017). This can be particularly useful in cases where the tumour is located near sensitive structures or organs that are located in areas of the body which move, such as the lung or liver, as it enables the radiotherapy to be delivered more precisely. MRI-based dosimetry involves using MRI to measure the dose of radiation that is delivered to the tumour and surrounding tissues during radiotherapy. This can help to ensure that the desired dose of radiation is delivered, while minimizing exposure to healthy tissue. Overall, the use of MRI in conjunction with radiotherapy can help to improve the effectiveness and safety of the treatment and reduce possible side effects.

1.4 Electron density information in PET and RT

Electron density information is important in external beam radiotherapy because it determines how much radiation is absorbed by different tissues in the body. To deliver a precise and effective dose of radiation to the target, it is important to know the electron density of the tissue in the area being treated. This is because the density of the tissue affects the amount of radiation that is absorbed and the depth at which the radiation will penetrate. For example, denser tissues, such as bone and muscle, will absorb more radiation than less dense tissues, such as fat or air. Knowing the electron density of the tissue can help the medical team to adjust the dose and intensity of the radiation to ensure that the tumour receives the appropriate amount of radiation while minimizing the risk of side effects to healthy tissues.

Similarly, AC of the bone is of extreme importance for the quantitative accuracy of PET, especially in the head region, where the bone surrounds the brain from all sides. The AC can be done in several ways. Utilizing a positron source, for example ^{68}Ge that electron captures into positron emitting ^{68}Ga, rotating around the subject on a rod, is called transmission-based attenuation correction (TXAC). TXAC is considered as a “gold standard” method for AC, as it allows the measurement of the gamma attenuation coefficients for each line of response directly.

Using positron source in AC derivation gives accurate AC, as the attenuation coefficients μ depend on the material and the energy of the photons in equation

$$I_x = I_0 e^{-\mu x} \tag{16}$$

where I_0 is the intensity at source, I_x is the intensity after the photons have passed through thickness of x (Bailey *et al.*, 2005, ch 2). For the equation to hold in PET, these attenuation coefficients μ need to be measured for each LOR with the same 511 keV energy gamma rays as the PET scan itself generates.

CT energies, however, are much lower, typically closer to the 80-150 keV range. As the attenuation coefficients μ increase while the photon energies decrease, a conversion from the attenuation coefficients measured with CT energies needs to be performed in order for them to be usable in PET. This causes the CT-based attenuation correction (CTAC) to be considered as a "silver standard" (Hitz *et al.*, 2014).

In CT, the attenuation is typically measured in Hounsfield Units (HU) where water is given the value of 0 HU (Hounsfield, 1973). Although Hounsfield initially set air to be -500, it is usually set to be -1000 HU. As water is 0 HU and air is -1000 HU at any X-ray photon energy, the HU values of other tissues change depending on the photon energy. This is because the ratio of attenuation coefficients μ of two different density materials changes when the photon energy changes (Hubbell, 1969). More precisely

$$\frac{\mu_{d,h}}{\mu_{s,h}} > \frac{\mu_{d,l}}{\mu_{s,l}} \tag{17}$$

where d and s denote dense and soft materials respectively, and h and l correspond to higher and lower photon energies. Yet, when the attenuation coefficients are close to each other, like in human tissues, the inequality in Eq. 15 is close to being an equality. Another source of error is the scattering of the X-ray photons that is different than the scattering of 511 keV annihilation photons due to the lower energy of the X-ray photons. These inaccuracies, contribute to the "silver standard" classification of CTAC.

When the HU values and CT photon energies are known, it is possible to convert the HU values into 511 keV attenuation coefficients μ. Several conversion methods have been presented in literature, some optimized for different tube currents than others (Kinahan *et al.*, 1998; Burger *et al.*, 2002; Bai *et al.*, 2003; Carney *et al.*, 2006; Ay *et al.*, 2011; Abella *et al.*, 2012), so the application of the attenuation coefficient conversion must be done carefully. Once the attenuation coefficients for each point along a LOR have been converted from HU values, the attenuation coefficient for the whole LOR can be calculated by integrating over the LOR

$$\mu_z = \int_0^y \mu(y)dy \tag{18}$$

where $\mu(y)$ is the attenuation coefficient μ at point y along the LOR z. Now finally, a measured incident count C_z along LOR z can be attenuation corrected using Eq. 14 as a starting point

$$C_{z_0} = \frac{C_z}{e^{-\mu_z}}. \tag{19}$$

Choosing a proper HU to attenuation coefficient conversion method is critical for the quality of the CTAC. However, as these are well established, and as TXAC is becoming unavailable, CT can be used as a reference AC standard for PET imaging to which other modalities can be compared. Similarly, for RT treatment plans (RTP) the dose delivered to the clinical target area depends on the initial intensity of the radiation and the density of the tissues between the skin and the tumour. Being certain about the tissue density is important, so that the dose does not need to be increased in order to counter uncertainties, instead only the minimal dose is needed for the patient. In this study, the focus is on MRI based AC (MRAC), which seeks to convert the MRI images into attenuation maps, or more generally expressed as substitute, or synthetic CTs (sCT), which can be applied to MRAC and MR-RT alike.

Using MRI instead of CT for PET imaging or for RT treatment planning is beneficial as it has better soft tissue contrast and no radiation dosage to the patient. On the downside, MRI images do not separate air from bone as well as CT images do, and this is problematic for PET or RT, especially in the head where the tissues of interest are typically located within the skull. The bone segment contributes strongly to the total, as it is the densest part of the head, but the accurate delineation of bone is difficult with conventional T1/T2-weighted MRI sequences. This is due to the short T2* relaxation time for bone. Various non-bone tissues have T2* relaxation times between 20 and 150ms (Bojorquez *et al.*, 2017), but bone T2* relaxation time is as low as 1–3 ms (Mahar *et al.*, 2018), which makes it difficult to distinguish from air. The resulting signal from bone is thus practically a void in the final MRI image. This problem arises from the differing physical phenomena captured by CT and MRI imaging. MRI images reflect the tissue relaxation and proton density whereas CT measures the X-ray attenuation which is related to the electron densities of the tissues. Moreover, the information MRI gathers is not simple to translate into electron density values. So, a model is needed to convert the MRI image into a sCT, which can then be used in the PET or RT pipeline.

The conversion of MRI images to HU values or gamma attenuation coefficients is not a trivial endeavour and has resulted in extensive research and development of sCT methods for PET-MRI and MR-RT. Some important applications of the sCT

conversion models are neurological PET-MRI imaging and MR-RT of the head region. Because the quantitative accuracy of combined PET-MRI and the accuracy of treatment plans of MR-RT are affected by accuracy and quality of the sCT used for AC or RTP, these applications require highly accurate methodology.

Several review articles have covered the AC in PET using a variety of different methods including methods for brain and whole-body imaging (Martinez-Möller and Nekolla, 2012; Berker and Li, 2016; Mehranian, Arabi and Zaidi, 2016a; Johnstone *et al.*, 2018; Teuho *et al.*, 2020). Recently, MRAC methods have been introduced, which achieve ±5% quantitative accuracy for PET images and less than 0.2% dose differences for RTP when compared with the CT-based approach (Andreasen *et al.*, 2015; Ladefoged *et al.*, 2017; Kemppainen *et al.*, 2019; Yu *et al.*, 2021). A rough division of the methods into three are as followed: atlas-, emission-, and segmentation-based methods (Martinez-Möller and Nekolla, 2012; Berker and Li, 2016; Mehranian, Arabi and Zaidi, 2016a; Johnstone *et al.*, 2018; Teuho *et al.*, 2020). To summarize, the atlas- and template-based methods enable the calculation of sCT by co-registering a single CT image or a set of CT images to the individual anatomy space of a subject's MRI image (Burgos *et al.*, 2014, 2015; Demol *et al.*, 2016; Mérida *et al.*, 2017). Moreover, several machine learning-based methods have been presented recently (Mecheter *et al.*, 2020), and deep learning has gained a lot of attention in sCT derivation (Krokos *et al.*, 2023). It is possible to use raw PET data to reconstruct the attenuation sinogram mathematically (Nuyts *et al.*, 1999; Rezaei *et al.*, 2012; Berker and Li, 2016; Mehranian, Zaidi and Reader, 2017). The PET-MRI systems capable of TOF may gain further improvements in quantitative accuracy (Conti, 2011; Mehranian and Zaidi, 2015b). There are also voxel-based methods that can be used to calculate the MRI intensity values of the voxel regardless of its spatial location, and hybrid methods are a combination of the voxel and atlas-based methods (Edmund and Nyholm, 2017; Johnstone *et al.*, 2018). Additionally, Bayesian probabilistic approaches, for head region bone modelling (Gudur *et al.*, 2014) and curve fitting with exponential and polynomial distributions for pelvic and head volumes (Kapanen and Tenhunen, 2013; Koivula, Wee and Korhonen, 2016) have been proposed.

In segmentation-based methods, the MRI images are segmented into different tissue classes, and the attenuation correction is built on top of said segmented tissue classes. Accurate bone delineation becomes challenging when conventional T1/T2-weighted MRI sequences are used, as the T2* relaxation time of bone is almost the same as air. Certain advanced methods for segmentation exist, such as probabilistic atlas-based segmentation (Chen *et al.*, 2017), which help to alleviate this issue. Advances in segmentation-based methods have shown a good level of accuracy when they have been compared to atlas- and emission-based methods (Ladefoged *et al.*, 2017). Outside the conventional MRI sequences, there are options, such as ultra-

short echo time (UTE) sequences and zero echo time (ZTE) sequences, which enable the visualization and segmentation of the bone directly from MRI images (Aasheim *et al.*, 2015; Delso *et al.*, 2015).

1.5 Machine learning methods

Machine learning (ML) is a subfield of artificial intelligence that uses algorithms to automatically improve a learning model and make predictions without having received explicit instructions on how to do so (Jordan and Mitchell, 2015). ML can be used to automate and improve the accuracy of image segmentation in MRI. One approach is to use supervised learning, where the algorithm is trained on a dataset of labelled MRI images, with each pixel in the image assigned to a specific tissue type or structure. The algorithm learns to identify patterns and features in the images that are characteristic of different tissue types and uses this knowledge to classify each pixel in a new image.

Another approach is to use unsupervised learning, where the algorithm is not provided with any labels or information about the tissue types in the images. Instead, it must learn to identify and segment different structures in the images based on their statistical properties and relationships with other pixels.

Both supervised and unsupervised machine learning approaches have been applied to MRI image segmentation with promising results (Mecheter *et al.*, 2020).The choice of approach and specific ML algorithms will depend on the characteristics of the MRI data, the desired level of accuracy, and the computational resources available. Several review articles focusing only on the ML assisted AC or imaging modality conversions have been written recently (Chen and Liu, 2023; Confidence *et al.*, 2023; McNaughton *et al.*, 2023).

Most studies have used full image data from which the ML algorithms find appropriate features and structures which then are regenerated for independent test images. Using whole images permits the usage of accurate deep learning neural networks. Deep learning networks are self-learning structures that can learn high level image features and model nonlinear connections between different image spaces (Lecun, Bengio and Hinton, 2015; Mecheter *et al.*, 2020). Training these models can take a long time, but prediction times for independent images are quite fast. Some studies with lower level data, such as voxel level conversion, exist where the images are sampled for smaller patches of data to be converted, rather than using the full image directly (Valverde *et al.*, 2017; Kläser *et al.*, 2021; Martinez-Girones *et al.*, 2021).

In a random forest (RF) algorithm, multiple decision trees are trained on random subsets of the data, and the final prediction is made by taking the average prediction of all the individual trees. This can be done by taking the mean of the predictions, or

by voting, where each tree “votes” for a particular class and the class with the most votes is the final prediction (Breiman, 2001). Mathematically, a decision tree is a flowchart-like tree structure that makes decisions based on a series of binary splits. For example, in a decision tree for classifying animals, the tree might first split on the feature “has wings” with a “yes” branch leading to another split for the feature “can fly” and a “no” branch leading to another split for the feature “has fins”. This process continues until the tree reaches a leaf node, which represents a prediction.

A distributed machine learning algorithm involves training many instances of the model in parallel and combining the learning process to help the final model converge faster (Zhou *et al.*, 2017). Each sub model can be trained on a random subset of the data, which can alleviate data transfer problems, in case different computers are used.

In a distributed random forest (DRF) algorithm, the training of the random forest is distributed across multiple processor cores, or computers, which can help to reduce the time it takes to train the model and make it more scalable. The training process is parallelized with each machine training a separate decision tree on a different subset of the data. The final prediction is made by aggregating the predictions of all the individual trees, just as in a regular random forest.

The use of multiple decision trees and the random sampling of the data help to reduce overfitting. Meaning that a model fits the training data too well and does not generalize well to new data. When multiple decision trees are trained on different subsets of the data, the random forest can capture a wider range of patterns in the data and make more accurate predictions.

A derivate of the DRF algorithm is called Extremely Randomized Trees (XRT) (Geurts, Ernst and Wehenkel, 2006). The novelty of the XRT algorithm is that the entire training dataset is sampled, and a tree is built for each observation, with varying set of features per tree. Furthermore, all splits are completely randomized, which should lower the variance and improve the computation time, as less time is needed to find the proper splits. The XRT method is also less susceptible to patterns in the training dataset, as no feature can be overrepresented.

1.6 Summary

Medical imaging is a tool to look inside a patient and learn about their anatomy and physiology. Scans done with some modalities like X-rays, and by extension CT, are somewhat straightforward in capturing images, while other modalities like MRI and PET require additional steps and need more complex mathematics in the process towards an accurate image result. MRI works by utilizing magnetic fields and radiofrequency signals to create detailed images of the body. PET on the other hand relies on radioactivity sources injected into the patient, so attenuation correction of

the measurement is essential. Tissue density estimation is also necessary for RT, so that the dose delivered to the target tissue is as accurate as possible. The tissue density estimation can be done in many ways, and in this study the focus is on MRI based methods and improving them.

MRI has superior soft tissue contrast and doesn't involve radiation exposure, and thus offers significant upsides for PET attenuation correction and electron density estimation for RT planning. However, distinguishing between different tissue densities, particularly bone and air, can be difficult for MRI. Conventional T1/T2-weighted MRI sequences struggle to accurately delineate bone due to the short T2* relaxation time of bone, resulting in signal from the bone tissue being practically absent and inseparable from air in the final MRI image. Thus, models are needed to accurately convert MRI images into sCT images, which can then be used in PET or RT. The conversion of MRI images to HU or gamma attenuation coefficients is a complex task that has led to extensive research and development of sCT methods for PET-MRI and MR-RT. Achieving high accuracy and quality in sCT methodology is crucial for applications, such as neurological PET-MRI imaging and MR-RT of the head region. Lately, there has been efforts to utilize machine learning to assist with the creation of sCT images. In this study, the focus is on improving the sCT creation in the head by setting up a pipeline to create sCT images from MRI images, then inspecting the effect of TOF in reconstruction error. Next, a novel way to automatically isolate and handle sinus regions in MRI images is developed, and finally the skull bone handling is improved with aid of machine learning.

2 Aims

The purpose of the study was to generate an offline attenuation correction pipeline for MRI images. Additionally, once the pipeline was established, the goal was to explore novel methods to improve different components of the pipeline. The study was split into four publications. The first task was to set the framework of the original pipeline, and the later studies focused on one aspect to improve or evaluated the performance of the initial pipeline.

2.1 Publication I – Initial pipeline framework development

In publication I (Teuho *et al.*, 2016) the aim was to set up a robust pipeline for attenuation correction using clinical anatomical MRI imaging sequences. The pipeline should work with minimal manual assistance, relying on automation in all phases of the process.

2.2 Publication II – TOF effect on reconstruction quality using different AC maps

Publication II (Lindén, Teuho, Klén, *et al.*, 2022) aimed to explore the effects of Time-of-Flight reconstruction using a variety of attenuation maps generated with the initial pipeline framework. The goal was to see how much time of flight improves the PET reconstruction accuracy even if the attenuation maps themselves had some inaccuracies.

2.3 Publication III – Sinus segmentation method development

The aim of publication III (Lindén, Teuho, Teräs, *et al.*, 2022) was to find a way to automatically address the problem of segmentation in the sinuses. The goal was to present new methods to model the air/soft tissue mix of the sinus region accurately, and to not have overrepresentation of bone tissue in the region.

2.4 Publication IV – Machine learning assisted bone segmentation and HU assignment

Publication IV aimed to replace the bulk probability assisted segmentation with a more accurate machine learning based approach. The goal was to use an RF algorithm to create an alternative bone probability to the one used previously in the sCT pipeline to assign the bone tissue more accurately, and test whether this output could be used to estimate the attenuation coefficient directly. Aim was also to use the machine learning on the voxel level to avoid artefacts of atlas-based conversion, and make sure even the smallest details of the subject's anatomy were converted with high accuracy.

3 Materials and Methods

3.1 Summary

This section consists of a brief overall summary of the methodology in each publication. Some aspects of the study are defined in more detail in separate sections.

3.1.1 Initial pipeline framework development

The process of creating the attenuation maps starts by segmenting the MRI images gathered with 3D T1-weighted fast field echo sequence into six tissue class probability maps using the New Segment function of SPM8 in MATLAB. The probability maps for grey matter, white matter, and cerebrospinal fluid were ignored, and only the probability maps for bone, air and soft tissues were used for the rest of the pipeline. The probability maps were converted into binary masks using 0.25 probability value as a threshold for each voxel. The soft tissue map was morphologically closed, and the closed mask was then used to create a mask for overall head shape. The bone mask was cleaned by finding the largest connected structure within the bone mask by connected component analysis. After some Gaussian filtering to remove any residual noise, the tissue masks were ready to be merged into one attenuation map.

To create the attenuation map, first the soft tissue volume is assigned an attenuation factor of 0.096 cm^{-1} (Catana *et al.*, 2010; Burgos *et al.*, 2014). All voxels belonging to the skull mask were assigned a bone attenuation coefficient of either 0.135 cm^{-1} or 0.145 cm^{-1} (Hofmann *et al.*, 2009; Catana *et al.*, 2010; Wagenknecht *et al.*, 2013; Anazodo *et al.*, 2015). The performance of different bone attenuation values was compared to the CTAC attenuation map. Tissue classes were finished by assigning all voxels that belong to the air mask with an attenuation coefficient of 0.0 cm^{-1} (Catana *et al.*, 2010; Burgos *et al.*, 2014). Lastly, the combined μ-map was smoothed with 5mm FWHM Gaussian filtering (Schramm *et al.*, 2013).

Tissue probability-based AC (TPB-AC) and CTAC μ-maps were then co-registered and resliced to the MRI image space. The PET reconstructions using TPB-AC and CTAC μ-maps and clinical MRAC μ-map were done with LOR-RAMLA

algorithm with 3 iterations and 33 subsets, with 128×128×90 image matrix with a voxel size of 2×2×2 mm. The performance of each model was evaluated by visual inspection using volume of interest (VOI) analysis with mean and standard deviation calculation, as well as correlation analysis. Additionally, a 28-region automatic anatomical labelling atlas (AAL) atlas was fit to each subject's images for brain region specific evaluations.

3.1.2 TOF effect on reconstruction quality using different AC maps

Three different attenuation maps were created and tested against CTAC attenuation maps. The first attenuation map, three-class model, was done using the same pipeline as earlier, with the soft tissue attenuation coefficient set to 0.096 cm^{-1} and the bone attenuation coefficient set to 0.151 cm^{-1}, which was one of the values tested in the initial pipeline setup. The same 3D T1-weighted fast field echo sequence was used as previously. The second map, two-class model, had a uniform attenuation coefficient of 0.100 cm^{-1} assigned to all soft tissue and bone voxels of the three-class model attenuation map. The third map, two-class model without air, had all air cavities inside the head also assigned with the attenuation coefficient of 0.100 cm^{-1}. The reference CTAC maps were created from a low-dose CT acquisition from which the HU values were transformed into linear attenuation coefficients using a bi-linear transformation (Burger *et al.*, 2002).

Both CTAC and MRAC maps were co-registered to non-attenuation corrected PET (NAC-PET) image space using SPM8 before the PET image reconstruction. The non-TOF PET and the TOF reconstructions were done with a line-of-response row expectation maximum algorithm (LOR-RAMLA), and ordered subset expectation maximum (TOF-OSEM) algorithm respectively. The results of the reconstructions were inspected visually and quantitatively. A quantitative analysis was done by using 35 cortical region VOIs from the AAL, after which the absolute differences between CTAC and MRAC reconstructed PET images were calculated for each VOI. The calculation was repeated for every MRAC method with the following equation:

$$\Delta\% = \frac{PET_{MRAC} - PET_{CTAC}}{PET_{CTAC}} * 100\ \%. \tag{20}$$

The images were also normalized by transforming them to the Montreal Neurological Institute (MNI) image space using SPM8. This was done to enable the calculation of the mean images.

3.1.3 Sinus segmentation method development

The initial pipeline from Publication I was used with some modifications, to see if the sinus region could be modelled better with the given MRI sequence. The pipeline was migrated to SPM12 from SPM8, and brain tissue was given its own tissue class and attenuation coefficient of 0.0985 cm^{-1}. The earlier tissue probability maps (TPM) based attenuation coefficient assignment was compared with a novel cuboid method and a template method, both of which use continuous attenuation coefficient assignment for the sinus region. CTAC maps were used as reference.

The cuboid method attempts to fit a cuboid shaped mask to the individual anatomy of each subject by varying the size and shape of the cuboid and choosing the cuboid with the highest cuboid goodness score. In the template method, a cuboid shaped mask is placed into the sinus region of an image in the MNI space. Then that mask is converted into each subject's individual space using the deformation vector fields given by the SPM normalization algorithm.

Once a sinus mask is defined, all the voxels within the mask are converted to a mixture of air and soft tissue based on the MRI intensities of the voxel. A stepwise function is used for the conversion. To perform the PET reconstruction, the CTAC and MRAC images are first co-registered to NAC-PET images, after which the reconstruction is done using the co-registered attenuation maps. Both non-TOF and TOF reconstructions were done to also assess the differences between reconstruction types for each sinus model.

The performance of the different models was evaluated with DICE analysis, bone region voxel correlation analysis, and sinus region attenuation coefficient analysis. A quantitative analysis was done by using the same AAL atlas as in the TOF/non-TOF evaluation and calculating VOI level differences. Finally, bias atlases were computed for visual inspection.

3.1.4 Machine learning assisted bone segmentation and HU assignment

The initial pipeline was used as a starting point for the ML method evaluation with minor modifications. The pipeline consisted of six tissue classes, each with their own attenuation coefficient. As the analysis for the ML method evaluation was done in HU scale, the bulk density attenuation coefficients were updated to have 942 HU for bone, 41 HU for grey matter, 25 HU for white matter, 15 HU for cerebrospinal fluid, 0 HU for soft tissue, and -1000 HU for air. All processing was done in SPM12.

A RF ML model was created with 8 subject training dataset and tested with a 10 subject test dataset. The quality of the conversion was evaluated with DICE coefficients, F-scores, and receiver operating characteristics (ROC) related statistics

such as area under curve (AUC) and Youden's J (Youden, 1950). The sinus conversion method based on sinus cuboids and MRI intensities was used for the sinus region. The RF probability was also used to estimate the bone HU values directly with a simple regression curve. This analysis was done with the leave-one-out method to utilize the limited sample size to maximum effectiveness.

3.2 Datasets

In total there were three subject datasets used in the study. The data for these studies was attained under the research permissions of Dnro T11/006/21, approval date: 19 January 2021 and Dnro 116/1801/2017, approval date: 21 Nov 2017.

3.2.1 Initial pipeline framework development

The dataset for the initial pipeline development consisted of PET CT and MRI data of seven subjects who were suspected of having memory disorders. The median and range of age, dosage, and weight were 47 (26–74) years, 274 (199–299) MBq, and 77 (47–80) kg, respectively. All subjects underwent a clinical PET-CT routine with either the Discovery VCT PET-CT (General Electric Healthcare, Milwaukee, US) or Discovery 690 PET-CT (General Electric Healthcare, Milwaukee, US). In addition to the PET-CT scans, PET and MRI acquisitions were performed with the Philips Ingenuity TF PET-MR.

3.2.2 TOF effect and sinus segmentation methods

Both TOF effect study and sinus segmentation method evaluation used the same dataset of subjects. The data consisted of 10 subjects (seven females, three males) who were suspected of having memory disorders. All subjects had undergone PET-MRI and PET-CT scans, and their data was used retrospectively. Seven subjects had their PET-MRI and PET-CT performed during the same session, and three on different days. The mean and standard deviation of the subjects age, dose ([^{18}F]-FDG), and weight were 53 ± 14 years, 277 ± 47 MBq, and 73 ± 19 kg. For the subjects who had undergone both PET-CT and PET-MRI sessions in the same day, the mean and standard deviations of the scan start times were 80 ± 20 min after the injection.

3.2.3 Machine learning assisted bone segmentation and HU assignment

The subject dataset was changed for the ML method evaluation. The dataset consisted of 18 RT subjects, nine of whom had brain metastases and nine had glioma. Ten were male and eight were female. The mean and standard deviation of age was 66.0 ± 12.5. Same dataset with slight changes was used in an earlier study (Ranta *et al.*, 2020).

3.3 Image segmentation

An MRI image segmentation algorithm divides the MRI images into several TPMs. Every map contains a voxel level probability to the tissue that the probability map corresponds to the algorithm used in this study was the New Segment function in SPM8 and SPM12 (Wellcome Trust Centre for Neuroimaging, University College London, UK). The pipeline setup, and TOF/non-TOF comparison, were done on SPM8 and the sinus, while the ML method evaluations were done with SPM12. The New Segment function combines affine registration to the International Consortium for Brain Mapping (ICBM) Tissue Probabilistic Atlas and image segmentation was based on the Unified Segmentation principle described in (Ashburner and Friston, 2005). The result of the MRI image segmentation is a series of TPMs, where a voxel corresponds to a probability between 0 and 1 of a particular tissue class to exist (Ashburner and Friston, 2005). Probabilities are scaled between the TPMs so that the total sum of tissue probabilities in a voxel over all TPMs is always one (Ashburner and Friston, 2005). The TPMs given by the New Segment function correspond to grey matter (GM), white matter (WM), cerebrospinal fluid (CSF), scalp, bone, and air.

3.4 Reconstruction pipeline for sCT

To initiate the sCT reconstruction, the DICOM data were exported to an off-line computer from PACS after which Nifti conversion was performed using DICOM import in SPM8 or SPM12. The MRI images were then segmented with the New Segment function in SPM. The New Segment function segmented the MRI images into six TPMs, which are WM, GM, CSF, scalp, bone, and air. For the sCT reconstruction in the initial pipeline, the TPMs for WM, GM, and CSF were not processed further, and the probability maps of the scalp, bone and air were used in the pipeline with MATLAB (2011a-2021b depending on the method). For the sinus and ML method evaluation, as well as the ignored three tissue classes, were processed further. Logical binary masks were created from these TPMs.

The scalp binary mask was the starting point to assign the total soft tissue volume. This was done with a slice-wise morphological closing and flood-fill operation. The result of this was a head volume mask that had the value 1 inside the head and the value 0 on the outside of the head.

A connected component analysis with 3D 6-voxel regional connectivity analysis was then performed on the skull binary mask to find the largest connected structure, which was then assigned as the final skull volume. This final skull volume is again a binary mask, like the soft tissue mask earlier. Residual noise was reduced with a Gaussian filtering, after which the filtered masks were converted back to logical masks with a threshold value of 0.01. Ultimately, the determined μ-values were given to each tissue type when the masks were combined to form the final three-class sCT. For the sinus and ML method evaluation, the soft tissue part of the sCT was further refined by giving separate attenuation coefficients for the other brain tissue types. Additionally, the volume within the sinus mask was then reconverted and given a continuous attenuation coefficient based on the sinus conversion model.

3.5 Time-of-Flight and PET reconstruction

The Philips Ingenuity TF PET-MRI (Philips Healthcare, Cleveland, OH, USA) scanner was used in all PET studies, and can reach the time resolution of 525 ps. In this study, the TOF reconstructions were done using BLOB-OSEM-TF reconstruction algorithm with blob and reconstruction parameters of alpha = 4.1338, radius = 2.3, blob increment = 2.0375, and relaxation parameter = 1.00.

3.6 Attenuation correction in sinus area

The sinus area attenuation correction method was derived using an anatomical MRI T1-weighted 3D fast field echo sequence, and it was later successfully applied in the ML bone segmentation model pipeline with T1-weighted 3D mDIXON sequence with gadolinium contrast agent.

To start the attenuation correction in the sinuses, an initial binary mask of the air inside the sinus cavity was created by taking the intersection of the air mask and the soft tissue mask (Figure 3). The mask was called inside air mask. This mask consisted of voxels that the segmentation considered to be partly air and partly soft tissue. This included the throat, large air cavities in the sinus, and noise outside the anatomical air cavities. To find the location of the largest air cavities, slice sums of the inside air mask were created and plotted. Based on comparisons of images and the slice sum plots, it was determined that most commonly the sinus regions were the local maximum of the slice sums right before the largest drop in the slice sums.

This is intuitive since above the sinuses there is not much air in the head. The slice after which the largest drop occurred was determined to be the slice i that maximises the formula

$$D = S_i - S_{i+d} \tag{21}$$

where D denotes the size of the drop, S_i is the slice sum of slice i and d was the length of the step. Values $d > 1$ allow for some noise to be present in the slice sum plot, as sometimes the largest drop did not occur just between two consecutive slices. The length d was set to be 10% of the number of slices, so that images with different slice counts would be treated similarly. Once slice i was found, the local maximum was simply the slice in range $\{i - d, \ldots, i + d\}$ that had the largest slice sum. This slice would be named the sinus starting slice.

Once the sinus starting slice was found, the matching air mask for that slice was eroded by a disk with a radius of 2. This removes most of the noise within the slice. Centre-of-mass (COM) voxel was calculated for all remaining separate pocket of air. These COM voxels would mark the starting point for the sinus air cavity search.

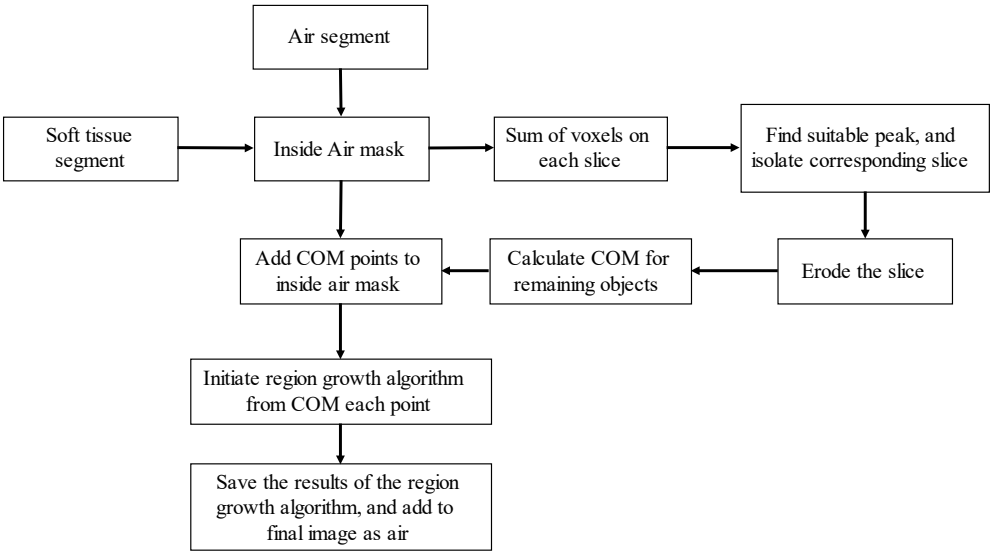

Figure 3. Air cavity algorithm process flow. The air and soft TPMs are the inputs, and the air cavity mask is the output. (Lindén, Teuho, Teräs, et al., 2022), sinus method evaluation.

3.6.1 Air cavity search

From each COM point, a recursive region growing algorithm was initiated across the initial inside air mask image. A new mask was created with just the COM points marked as air. A recursive algorithm called CavityCreeper was started from those COM points.

Summarised CavityCreeper in MATLAB:

```
function out_im=Creeper(ref_im, in_im, Coords, layer, maxd)
   if out_im(Coords)==1 && layer<maxd
         neighbour=out_im(newCoord);
         ref=ref_im(newCoord);
         if neighbour==0 && ref == 1
             out_im(newCoord)=1;
             out_im=Creeper(//
                       ref_im,out_im,newCoord,layer+1,maxd);
         end
    end
end
```

where `out_im`, `ref_im` and, `in_im` represent the output image, reference image, and input image respectively. `Coords` represent the initial COM point coordinates, and `newCoord` is always one of the three-dimensional 6-neighbourhood coordinates where exactly one of x, y, or z dimensions is altered by one. Although not depicted in in the summarized version of the algorithm above, each call of the recursion algorithm initiates six more recursive calls due to six possibilities for `newCoord`. Of course, some of those calls would not trigger more calls if they ended up in the previously handled voxels, or met only non-air voxels, and thus could not expand. The `maxd` parameter of 20 was added so that the recursion would eventually stop in case the recursion would end up in the air surrounding outside the head, or alternatively there was excessive noise in the image. The cavities found by the CavityCreeper would be set as air in the final image.

3.6.2 Cuboid and Template methods

To get a cuboid fitting the subject anatomically, a large set of cuboids were placed around the air cavities of the subject to find the cuboid with highest Cuboid Goodness (CG). Initially, a maximum size cuboid was automatically placed in the anterior part of the head, and the cuboid was set large enough to cover the air cavities. Following this, the possible sub-cuboids within the largest possible cuboid were placed into the image, and their CG was evaluated. The cuboid with largest CG was chosen for further processing, and there was a trade-off between the size of the cuboid and the amount of bone segment covered. The CG criterion is:

$$CG = \frac{\sum_{i \in C} p_i}{\left(\frac{c}{4}\right)^2}, \tag{22}$$

Where CG is the value of the cuboid goodness, C is the cuboid being tested, c is the perimeter of the cuboid, and p_i is the bone segment probability value for the voxel i. The CG value grows when voxels with values larger than 0 in the bone TPM are found within the tested cuboid. At the same time, growing the cuboid excessively is punished by the denominator that accounts for the perimeter of the cuboid. Perimeter was chosen instead of volume, as this favours more cube shaped cuboids. Whereas the volume denominator tended to create very flat cuboids, which did not cover enough of the bone TPM.

The second method for finding the sinus region was converting a cuboid template in the normalised space to an individual anatomy by using the inverse transformation fields created by the segmentation process. This method started by manually drawing a cuboid covering the sinus region in the MNI space, like the regional masks used in (Ladefoged CN *et al.*, 2015). As each subject was normalized with SPM12 during the segmentation, those inverse transformation fields were then used to transform the cuboid template back to the individual space of the subject. The mask that is generated by the inverse transform, is no longer a square cuboid, as different deformations are applied to the template because of the tilts and varying shapes present in individual subjects' images. The inverse transformed cuboid was used as a basis for further processing for the template method.

3.6.2.1 Cuboid search

The starting point of the cuboid search was the average of the COM points described earlier. The cuboid search would iteratively make a larger cuboid and vary its position around the starting point. For any image, the dimensions of it can be stated by using constants $x_{max}, y_{max}, z_{max} \geq 1$, one of each dimension. The size and location matrix for a cuboid within the image is a three-dimensional binary matrix $C = (C)_{i,j,k}$ with values 0 for most elements, and value 1 for the elements within cuboids that have indices in range

$$\begin{cases} i = x_s, x_s + 1, \dots, x_e \\ j = y_s, y_s + 1, \dots, y_e \\ k = z_s, z_s + 1, \dots, z_e \end{cases} \tag{23}$$

where start and end indices for each dimension are constrained by

$$\begin{cases} 1 \leq x_s \leq x_e \leq x_{max} \\ 1 \leq y_s \leq y_e \leq y_{max} \\ 1 \leq z_s \leq z_e \leq z_{max}. \end{cases} \tag{24}$$

There is also the bone probability image matrix P that consist of elements $p_{i,j,k}$ that denote the bone probabilities of voxels in coordinates i, j, k. Matrices P and C have the same dimensions. The goodness criterion for the cuboid can now be expanded as:

$$CG(C) = \frac{\sum_{i=x_s}^{x_e} \sum_{j=y_s}^{y_e} \sum_{k=z_s}^{z_e} p_{i,j,k}}{\left(\frac{d(C)}{4}\right)^2}, \tag{25}$$

where

$$d(C) = 4\big((x_e - x_s + 1) + (y_e - y_s + 1) + (z_e - z_s + 1)\big), \tag{26}$$

where CG is the value of the cuboid goodness, C is the cuboid being tested, $d(C)$ is the perimeter of the cuboid C, and $p_{i,j,k}$ is the bone segment probability value for the voxel i, j, k.

The square in the denominator further constrains the cuboid size, as it will grow fast when the perimeter grows. Perimeter instead of volume in the denominator favours more cube shaped cuboids than flat cuboids, since the cuboid volume for a given total perimeter is maximized with a cube. However, as the bone probabilities $p_{i,j,k}$ tend not to fall in an exact cube, the outcome cuboids were slightly flattened. Once all possible cuboids had been evaluated, the one with the highest CG was chosen as the final sinus cuboid. Based on image assessment, a rule was put in the algorithm that extended the final cuboid slightly forward and upwards, in order to capture the edges of the sinuses and nose.

3.6.2.2 Computationally efficient integral images

The computational demands for CG calculations are high, as the fitting area contains many possible cuboids, and each individual cuboid can contain thousands of voxels. To help alleviate these computational demands, a method to reduce the complexity was implemented. The method is similar to the formulae previously published in an earlier study (Tapia, 2011). The purpose of the improved method is to replace the repeated calls to probability matrix P with an iteratively grown sum matrix.

The calculation order is important to help the faster calculation process. Let the cuboid set Γ_M include all the cuboids $C_1, \dots, C_m$ that can fit inside the largest possible sinus cuboid C_m. A cuboid $C_\omega \in \Gamma_M$ can be defined by the coordinates of two of its opposite corners $x_\omega^s, x_\omega^e, y_\omega^s, y_\omega^e, z_\omega^s, z_\omega^e$, where x, y, z are the voxels corresponding to the dimensions in the cartesian coordinate system, s, e notate the points for the start, and the end corners of the cuboid respectively. For each $\omega \in \{1,2, \dots, m\}$, the following inequalities hold:

$$\begin{cases} x_m^s \le x_\omega^s \le x_\omega \le x_\omega^e \le x_m^e \\ y_m^s \le y_\omega^s \le y_\omega \le y_\omega^e \le y_m^e \\ z_m^s \le z_\omega^s \le z_\omega \le z_\omega^e \le z_m^e. \end{cases} \tag{27}$$

For the CG numerator, the sum σ_ω of all elements $p_{i,j,k} \in P$ that are located within the cuboid C_ω needs to be calculated.

$$\sigma_\omega = \sum_{i=x_\omega^s}^{x_\omega^e} \sum_{j=y_\omega^s}^{y_\omega^e} \sum_{k=z_\omega^s}^{z_\omega^e} p_{i,j,k}\,. \tag{28}$$

The three nested sums require many calls to matrix P. This corresponds to long data access and processing times when the number of cuboids m is large. It is also evident that the sums σ_ω share a lot of elements to $p_{i,j,k}$. Therefore, a large time savings can be achieved, if elements $p_{i,j,k}$ are not accessed multiple times.

Consider now matrix M_r, which consists of all the sums σ_r , that are calculated from cuboids $C_r \in \Gamma_r \subset \Gamma_M$. The set Γ_r is defined as the set of all the cuboids C_r, which share the starting corner x_r^s, y_r^s, z_r^s. Because of this shared starting corner, the set Γ_r is a strict subset of the set Γ_M, which has cuboids with varying start and end corners.

Each cuboid C_r has a different opposing end corner. The relative location of the opposite end corner of a cuboid C_r to the starting corner x_r^s, y_r^s, z_r^s will correspond to the location of the sum σ_r in the matrix M_r. The matrix M_r only covers the space of the largest cuboid within Γ_r and is thus smaller than matrix P that spans the whole image space. Due to the different sizes of matrices of M_r and P, a new sub matrix D_r is defined to ease the notation later. The matrix D_r has the same size as matrix M_r, and consists of elements of matrix P in a way that $D_r(1,1,1) = P(r_x, r_y, r_z)$, where the indices r denote the start corner of the cuboids C_r: $r_x = x_r^s, r_y = y_r^s, r_z = z_r^s$. Due to the simplicity of the first sum, it is trivial that $M_r(1,1,1) = P(r_x, r_y, r_z) = p_{r_x,r_y,r_z} = D_r(1,1,1)$.

If one of the indices is incremented by one, this includes one more element from D_r to the sum $M_r(2,1,1) = D_r(1,1,1) + D_r(2,1,1)$. Since it was already established that $M_r(1,1,1) = D_r(1,1,1)$, the previous equation can be written as $M_r(2,1,1) = M_r(1,1,1) + D_r(2,1,1)$. This recursive behaviour generalizes into one dimensional recursion equations along the i, j, k axes:

$$\begin{cases} M_r(i,1,1) = M_r(i-1,1,1) + D_r(i,1,1) \\ M_r(1,j,1) = M_r(1,j-1,1) + D_r(1,j,1) \\ M_r(1,1,k) = M_r(1,1,k-1) + D_r(1,1,k). \end{cases} \tag{29}$$

After the matrix M_r is populated with all of the elements along the axes, two dimensional equations can be reached. The sum $M_r(2,2,1)$ can now be calculated as $M_r(2,2,1) = D_r(1,1,1) + D_r(2,1,1) + D_r(1,2,1) + D_r(2,2,1)$, if term $D_r(1,1,1)$ is

added and subtracted to the previous equation, and terms from matrix M_r are substituted in, the following equation is reached

$$M_r(2,2,1) = M_r(2,1,1) + M_r(1,2,1) - M_r(1,1,1) + D_r(2,2,1). \quad (30)$$

Again, the recursive equation remains symmetric along any two-axis combination. Thus, the general two-dimensional recursion equations can be written as follows:

$$\begin{cases} M_r(i,j,1) = M_r(i-1,j,1) + M_r(i,j-1,1) \\ \qquad -M_r(i-1,j-1,1) + D_r(i,j,1) \\ M_r(i,1,k) = M_r(i-1,1,k) + M_r(i,1,k-1) \\ \qquad -M_r(i-1,1,k-1) + D_r(i,1,k) \\ M_r(1,j,k) = M_r(1,j-1,k) + M_r(1,j,k-1) \\ \qquad -M_r(1,j-1,k-1) + D_r(1,j,k). \end{cases} \quad (31)$$

After the two-dimensional elements to the matrix M_r have been calculated, the three-dimensional equations can be formulated. Consider now element $M_r(2,2,2)$. It can be calculated directly as a sum of elements in D_r as

$$\begin{aligned} M_r(2,2,2) = {} & D_r(1,1,1) + D_r(2,1,1) + D_r(1,2,1) + D_r(1,1,2) \\ & + D_r(2,2,1) + D_r(2,1,2) + D_r(1,2,2) + D_r(2,2,2) \end{aligned} \quad (32)$$

By adding and subtracting terms from matrix D_r and then substituting in suitable terms from the matrix M_r, the Equation 30 can be written as

$$\begin{aligned} M_r(2,2,2) = {} & M_r(1,2,2) + M_r(2,1,2) + M_r(2,2,1) - M_r(1,1,2) \\ & - M_r(1,1,2) - M_r(1,2,1) - M_r(2,1,1) + M_r(1,1,1) \\ & + D_r(2,2,2). \end{aligned} \quad (33)$$

This equation can be written in general form as:

$$\begin{aligned} M_r(i,j,k) = {} & M_r(i-1,j,k) + M_r(i,j-1,k) + M_r(i,j,k-1) \\ & - M_r(i-1,j-1,k) - M_r(i-1,j,k-1) \\ & - M_r(i,j-1,k-1) + M_r(i-1,j-1,k-1) \\ & + D_r(i,j,k). \end{aligned} \quad (34)$$

Note that the simpler one- and two-dimensional recursion equations can be derived from the last three-dimensional Equation 32, by setting a combination of the indices i,j,k to value 1 and then removing all terms from the three-dimensional equation that has any index with value of zero. To calculate an element in the matrix M_r, only up to seven calls needs to be made to matrix M_r and one to matrix D_r, compared to potentially thousands of calls to the matrix P in the earlier sum formula for σ_ω.

Albeit fast, there are still limitations on the order of calculation. Due to each new element requiring specific earlier elements to be calculated beforehand, parallel processing of these calculations is difficult, as computations need to be synchronised carefully. However, the already existing sum matrix M_r can be used to calculate sum matrixes M_o, which corresponds to the set of cuboids $C_o \in \Gamma_o \subset \Gamma_M$ that share a

starting corner x_o^s, y_o^s, z_o^s in a way that all the indices of the start corner are at least as large as the starting corner for the previously calculated set Γ_r

$$x_r^s \leq x_o^s, \qquad y_r^s \leq y_o^s, \qquad z_r^s \leq z_o^s. \tag{35}$$

If $x_r^s = x_o^s,\ y_r^s = y_o^s,\ z_r^s = z_o^s$, then trivially $M_r = M_o$. Therefore, it may be assumed that at least one of the inequalities is strict. If exactly one of the starting corner indices is different, then the element in matrix M_o can be calculated by subtracting the sum of the probability elements $p_{i,j,k}$ from the element in matrix M_r, that are outside the cuboid limited by the starting corner x_o^s, y_o^s, z_o^s. Practically, this can be done by utilizing index shift number $t = x_o^s - x_r^s$, or similar for y, z dimensions, to denote the part of matrix M_r that needs to be subtracted. The resulting equation for the subtraction is

$$M_o(i,j,k) = M_r(i+t,j,k) - M_r(t,j,k). \tag{36}$$

The equation is very similar if it is written along the other two axes:

$$\begin{cases} M_o(i,j,k) = M_r(i,j+u,k) - M_r(t,j,k) \\ M_o(i,j,k) = M_r(i,j,k+v) - M_r(t,j,k)' \end{cases} \tag{37}$$

where $u = y_o^s - y_r^s$ and $v = z_o^s - z_r^s$.

For the two-dimensional situation, similar subtraction happens, but due to the two subtracted elements from matrix M_r containing several same probability elements $p_{i,j,k}$, this intersect needs to be added back to the equation. The equation then becomes:

$$\begin{aligned} M_o(i,j,k) = {} & M_r(i+t,j+u,k) - M_r(i+t,u,k) - M_r(t,j+u,k) \\ & + M_r(t,u,k), \end{aligned} \tag{38}$$

and similar equations may be written for other two-axis combinations. Finally, in the three-dimensional case, where all three indices differ, the intersects of the three subtracted two-dimensional elements that are added back to the equation intersect with each other. This additional intersection term needs to be removed from the final formula. Thus, the final equation is as follows:

$$\begin{aligned} M_o(i,j,k) = {} & M_r(i+t,j+u,k+v) \\ & - M_r(i+t,j+u,v) - M_r(i+t,u,k+v) \\ & - M_r(t,j+u,k+v) + M_r(i+t,u,v) \\ & + M_r(t,j+u,v) + M_r(t,u,k+v) - M_r(t,u,v). \end{aligned} \tag{39}$$

Like earlier, the one- and two-dimensional equations can be reached from the last three-dimensional ones, by setting any of the index shift numbers t, u, v as zero and eliminating the terms where any of the indices go to zero.

If the set Γ_r from which the matrix M_r was calculated included the call to the cuboids that had the same starting corner as cuboid C_m, then the matrix M_r includes

the sum σ_m, which is the sum of all the elements in the largest possible cuboid C_m. Any cuboid $C_\omega \in \Gamma_M$ has a starting corner that is at least as large as the starting corner of the cuboid C_m, meaning $x_m^s \leq x_\omega^s,\ y_m^s \leq y_\omega^s,\ z_m^s \leq z_\omega^s$. This means that if the cuboid C_m belongs to the set Γ_r, then the all sum matrices M_o corresponding to all sets $\Gamma_o \subset \Gamma_M$ can be calculated from the sum matrix M_r. The calculation process for the sum matrices can start by first calculating the matrix M_r for a set Γ_r that contains the largest possible cuboid C_m. After this, every other sum matrix M_o can be calculated in parallel utilizing the already calculated matrix M_r. Once an arbitrary sum matrix M_ω pertaining to the set Γ_ω, is completed the CG value for all associated cuboids can be calculated simply by using the sum matrix in the simplified version of the Equation 23, CG formula

$$CG\left(\Gamma_{\omega_{i,j,k}}\right) = \frac{M_\omega(i,j,k)}{(i+j+k)^2}, \tag{40}$$

where i, j, k are the coordinates of the end corner relative to the starting corner. The denominator $\frac{d(C)}{4}$ is simplified, as all the starting points can now be set to value 1 as the indices i, j, k equal to the differences in the full version of the equation.

This time the parallelization can be mostly unsynchronized, as each set Γ_ω can keep calling the already calculated matrix M_r, and only needs to store the start and end corner values of the cuboid with highest CG value and the CG value itself. For example, a thread can be given to each set Γ_ω, and then compute the matrices M_ω sequentially to avoid I/O conflicts. Comparing the ultimate set outputs and finding the cuboid with the globally highest CG, once all the sets Γ_ω have been completed, is computationally fast. Unfortunately, true parallelization was not possible for the server setup used in the sinus method evaluation, so the backwards stepping part of the calculation process was not implemented.

3.6.3 Sinus area voxel conversion

When the cuboid location was found with either the sinus cuboid or cuboid template method, the voxels within it needed to be assigned HU values. This was done with an MRI-CT conversion model that converts the voxels into air, soft tissue, or a mix of the two. The data scatter plot (Figure 4) of the MRI intensity and CT HU values voxels within the cuboid masks showed a concentration of data at -1000 HU and 0 HU. Typically, the -1000 HU concentration point was applicable for the smaller MRI intensity values, whereas the 0 HU concentration point was present throughout the MRI intensity range. There is also a substantial number of voxels between -1000 HU and 0 HU concentration lines.

A tri-linear conversion function was created. The function first had a horizontal line matching air with -1000 HU until MRI intensity a. This was followed by a

diagonal line that starts at -1000 HU at MRI intensity a, and continued to 0 HU at MRI intensity b. This diagonal part represented the soft tissue – air mix within the sinuses. Lastly, another horizontal line at 0 HU started from MRI intensity b, and continued to maximum MRI intensity. The MRI intensities a and b that define the model were points were estimated from the data. The process was a bootstrap style sampling of 100 voxels from each subject's sinus mask, resulting in 1000 pairs of MRI intensities and HU values. All reasonable combinations of start and end values for the horizontal lines and the diagonal were tried. The one that gave rise to the smallest square sum of distances (SSD) of voxels to the conversion curve was chosen for that sample.

$$SSD = min \sum_{j=1}^{n} \begin{cases} \left(y_j - 1000\right)^2, if\ x_j < a \\ d\left(\left[x_j, y_j\right], y = \dfrac{1000}{b-a}(x-b)\right)^2, if\ a < x_j < b \\ y_j^2, if\ x_j > b \end{cases}, \qquad (41)$$

where a and b are the MRI intensities at start and end x-coordinates for the diagonal part, and parameters to adjust to minimize the *SSD*, and $d(x)$ is the distance formula for point $\left[x_j y_j\right]$ and the diagonal line segment in two-point form.

The sampling and estimation were repeated for 100 times, after which the average start and end points a and b were chosen as the final start and end points for the changes in the conversion function.

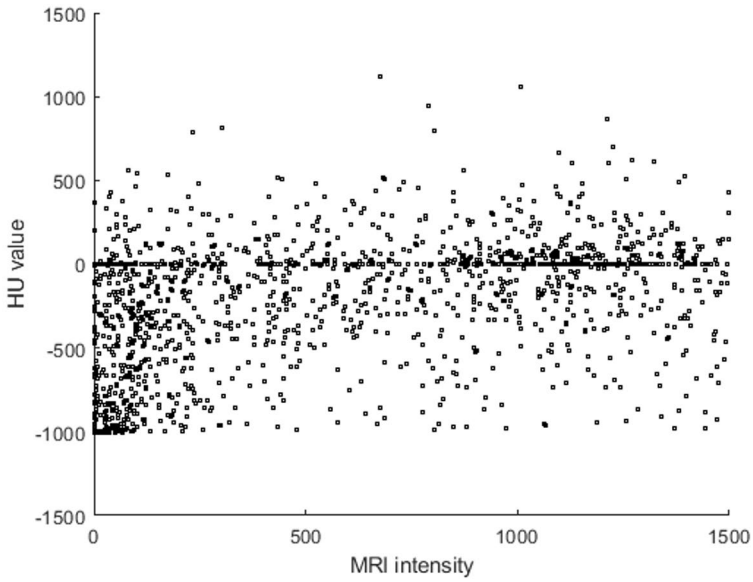

Figure 4. Summary of the data points of all MRI-CT sinus conversion samples. The voxels that are considered as bone based on their HU-values cannot be identified using their MRI intensity alone. (Lindén, Teuho, Teräs, et al., 2022), sinus method evaluation.

3.7 Machine learning methods

In the last part of the study, the goal was to improve the segmentation of the bone from the simple thresholding method originally implemented in the beginning. It was chosen to use machine learning methods to help with the segmentation. There were two MRI images available for each subject, called in-phase and out-phase images. These were both segmented into six tissue class probability maps, resulting in a total of 14 images for each subject.

It was also decided to not use an atlas-based approach to try to convert the whole image using the machine learning algorithms, but instead only gather data from the nearest environment of each voxel. This structured data approach supported a more limited cohort of test subjects to be used, as each subject can be sampled for millions of data points, if necessary. Ultimately, the dataset needs to be large enough to capture enough subject-to-subject variance to prevent overfitting and be more representative for independent test datasets.

Each subject was sampled for 30,000 voxels. Any voxels that were outside a simply drawn head mask were discarded, and new voxels were sampled until the set of 30,000 was reached. To provide the data some environment information, the average tissue probabilities, and MRI intensities from the 7×7×7 and 21×21×21 voxel environments of each data point were sampled. Together, with the sinus cuboid information, this leaves 43 data points for each voxel.

An RF machine learning algorithm, XRT, within open-source machine learning platform H2O was used. To begin, the 18-subject dataset was split into two groups. Set A had eight subjects, while set B had ten. The RF algorithm was used to classify whether a given voxel is bone or not by using 157 HU as the bone threshold. This was trained with Set A, and the resulting model was used to estimate the bone probabilities p_1 for Set B. Using a suitable p_1 threshold, removes most of the non-bone tissue, assigning better bone segmentation than the original one.

The bone probabilities p_1 and corresponding CT HU values in the Set B were also used to fit a simple 2nd degree polynomial regression curve between the RF model classifier probability and CT HU value:

$$sHU = a\,p_1^2 + b\,p_1 + c, \tag{42}$$

where sHU is the sCT HU value and p_1 is the RF classifier probability. The parameters a, b, c were chosen so that the sum of residuals is minimized:

$$\min\left(\sum (HU - sHU)^2\right), \tag{43}$$

where HU and sHU are the HU values for CT and sCT, respectively. This was done by leave-one-out method to generate 10 different conversion curves, one for each subject. The fitting was done similarly to an earlier study (Ladefoged CN *et al.*, 2015), and the HU values were averaged for each classifier probability (0, 0.01, 0.02, …, 0.99, 1.00) before fitting the regression curve.

3.8 Image quality assessment

The resulting images and distributions from all evaluations were assessed with a variety of different measurements, chosen for their applicability for each specific situation.

3.8.1 Descriptive tests

Descriptive tests were used to assess the overall agreement between images or distributions. These include arithmetic mean and median, standard deviation, and quartile range, as well as maximum / minimum values. Additionally, various graphs, such as scatter and Bland-Altman plots, were used to illustrate the differences between distributions. Mean average error (MAE) values (Eq. 42) were calculated as well when the ML based segmentation was evaluated

$$MAE = \frac{\sum_{i=1}^{n} |CT_i - sCT_i|}{n}, \tag{44}$$

where CT_i and sCT_i are the HU values for voxels within the inspected mask for CT and sCT images, respectively. The mask includes n voxels.

3.8.2 DICE coefficient

DICE coefficients were evaluated for the sCT-CT overlap in the sinus method and ML based segmentation evaluations to estimate the quality of the bone delineation in the attenuation maps. DICE coefficients for CT (A) and sCT (B) were calculated using Equation 43 (Dice, 1945):

$$DICE = \frac{2\,|(A \cap B)|}{|A| + |B|}, \tag{45}$$

Where the function |X| denotes the number of voxels in an arbitrary binary mask X. Voxels, where $HU > 157$, were included in the masks. A binary mask is an image which only get values of 0 and 1, or False and True, and are used to select a desired subsections of other images by multiplying the images with the binary masks. Binary masks can also be used to count voxels, as done above. DICE coefficient can vary

between 0 and 1, where coefficient of 0 means no overlap between the considered images, and 1 means perfect overlap or equivalently, that the intersection mask of the two images is the same as the union mask of those two images.

An HU threshold was fixed to 157 HU throughout the study when DICE coefficients were assessed. This is equivalent to the attenuation coefficient of $0.105\ cm^{-1}$, which is below the attenuation coefficient limit of spongy bone, reportedly to be at $0.11\ cm^{-1}$ (Mehranian, Arabi and Zaidi, 2016b). The attenuation coefficient is still higher than soft tissue, so it would separate the bone and non-bone tissues well. The choice to choose this value was supported by its previous use as a threshold between bone and non-bone tissue (Son *et al.*, 2010).

3.8.3 Attenuation coefficient and HU value analysis

The attenuation coefficient or HU values were investigated in every part of the study in various forms. Since improving the modelling of the attenuation maps was the main purpose of the study, the attenuation coefficients need to be investigated to properly assess the accuracy of the model. In the sinus model evaluation, the attenuation coefficients were assessed with a leave-one-out method, where the sinus conversion curve estimates were created without one of the ten subjects, and the resulting estimate was then used for the left-out subject. This was repeated with each subject.

3.8.4 VOI level analysis

As TOF effect and sinus model evaluations included in the PET data, it was reasonable to investigate the performance of the models on VOI level. VOI level analysis was generally done by examining the average values within the VOI using arithmetic means of the voxels within the VOI.

3.8.5 Bias Atlases

In the sinus model evaluation, the Atlas PET images were calculated. The Atlas images represented the mean and standard deviation of PET uptake (kBq/mL) across the subject group. The images were calculated for all evaluated CTAC and MRAC methods. To assess the global bias distribution, mean bias atlas images were calculated as described by (Ouyang *et al.*, 2013; Teuho *et al.*, 2017). Before inspection, the atlas images were masked with a binary brain mask.

3.8.6 ROC, AUC, and Youden's J

In the ML based segmentation evaluation, different classifiers were evaluated and compared against each other. ROC, AUC, and Youden's J were used to quantify the performance of the classifiers. ROC is a curve that connects the sensitivity and specificity of a classifier together. For each probability p_1 of the classifier both sensitivity and specificity are measured, and these pairs are plotted and connected for each value of $p_1 \in [0,1]$. A perfect classifier would have sensitivity that goes immediately to 1, and specificity that remains at 1 for all p_1. For plotting purposes, ROC is typically plotted with sensitivity on y-axis and 1-specificity on the x-axis. Thus, good classifier ROCs tend to get close to the top left corner, while bad ones remain close to the diagonal.

AUC refers to the area under the ROC curve. For perfect classifier AUC is 1, and for a perfectly random classifier, the AUC value is 0.5, as the ROC curve is a diagonal thereby making AUC effectively a triangle. Youden's J is the largest gap along the y-axis between the ROC and the diagonal line. The best possible Youden's J is 1, if the ROC jumps immediately to coordinates (0,1) while the diagonal is still at coordinates (0,0). A random classifier following the diagonal line would have Youden's J of 0.

4 Results

Chapter 4 summarizes the key results from the study. The initial pipeline development focuses on the choosing of appropriate attenuation coefficients for the sCT pipeline. The results indicate that the presented pipeline can generate usable sCT images for further applications. TOF effect evaluation explores the accuracy of PET reconstruction when the attenuation map is worsened, and how that accuracy changes when TOF or non-TOF reconstruction is applied. Generally, the TOF reconstruction reduces the error in the images as expected, and the error reduction is larger when the attenuation maps are of worse quality.

The sinus model evaluation assesses the performance of the new sinus methods by comparing DICE values, sinus area attenuation coefficients, and PET reconstruction accuracy between methods. The new sinus handling methods generate higher DICE coefficients and yield smaller errors in attenuation coefficient and show improved PET VOI level results with TOF reconstruction. The ML assisted segmentation assessment validates the quality of the ML RF model by DICE coefficients, MAE values, and classifier ROC performance metrics. The RF model shows improved DICE coefficients and reduced MAE values with continuous HU assignment. The RF classifier is also better when measured with the AUC and Youden's J.

4.1 Findings – Initial pipeline framework development

The SPM8 probability maps showed a reasonable segmentation, albeit some noise was present in the probability maps, and the sinus regions had too much volume in the bone TPM. An example set of anatomical MRI (atMRI) images can be seen in Figure 5. The resulting attenuation maps can be seen in Figure 6. Unlike the clinical attenuation map, the TPB-AC map had skull and air cavities, although the sinuses and mastoids had some issues.

The head level VOI analysis showed that the difference of mean μ-values ranged from -4.5% to +0.4% for the CTAC and TPB-AC when the attenuation coefficient of 0.145 cm^{-1} was used. There was more underestimation with 0.135 cm^{-1} as bone attenuation coefficient. The R^2 value for the whole head ranged

from 0.84 to 0.92 when the bone attenuation coefficient of 0.145 cm^{-1} was used. With the other tested attenuation coefficient (0.135 cm^{-1}) for bone, the results were slightly worse.

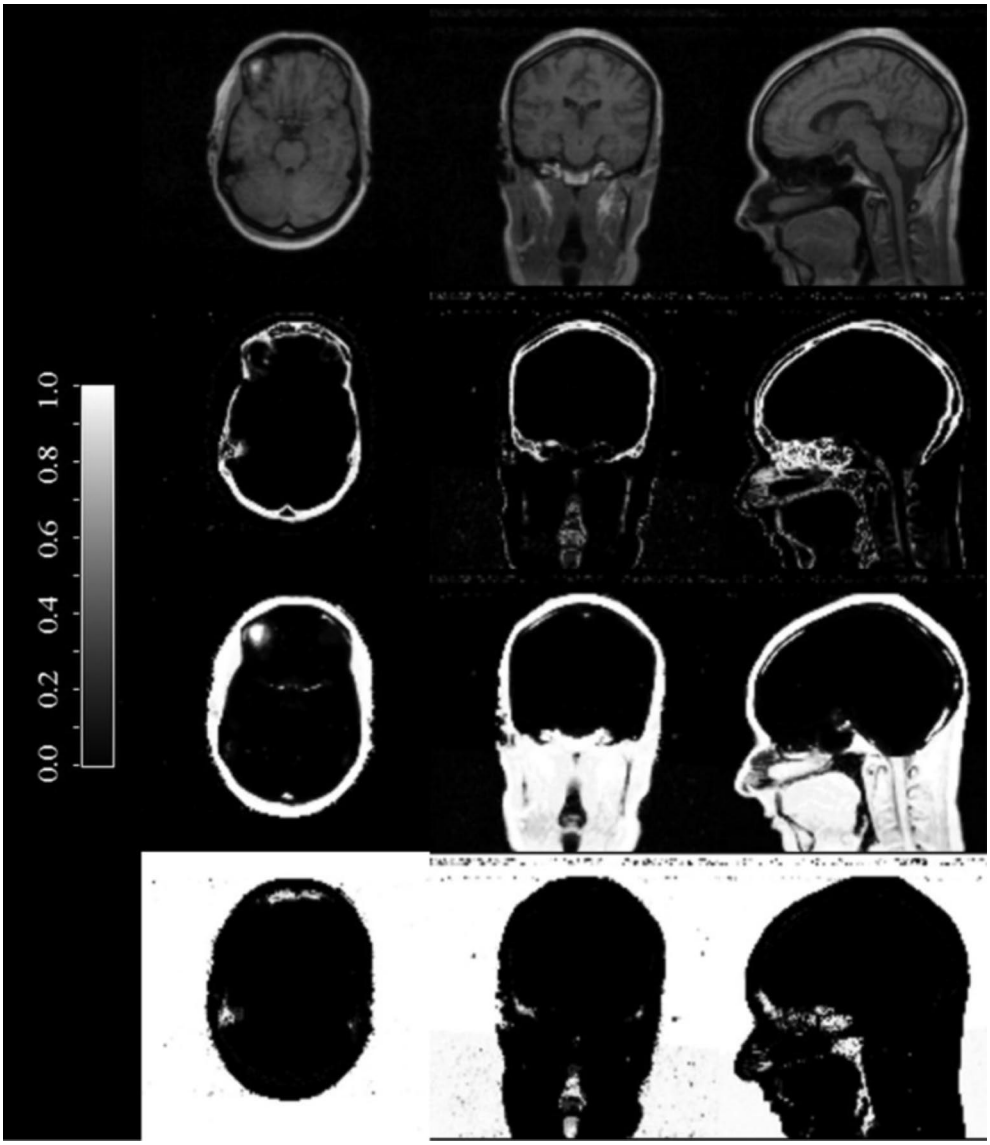

Figure 5. Example atMRI pictures, (Teuho et al., 2016), initial pipeline setup. Top row is anatomical MR data. Second, third, and fourth row is non-processed tissue probability maps for bone, scalp and air respectively. © 2016 IEEE.

There were smaller differences between the attenuation maps for the soft tissue parts of the image. The differences ranged from -1.4% to 2% when using a higher attenuation coefficient for bone. The subject-wise errors were larger in the skull VOI. The better performing three subjects had errors of -4.1%, -6.6% and -8.2%, while the errors of the others ranged from -9.8% to -11.6% when the bone attenuation coefficient of 0.145 cm^{-1} was used. Like earlier, the underestimation was more severe with a lower attenuation coefficient.

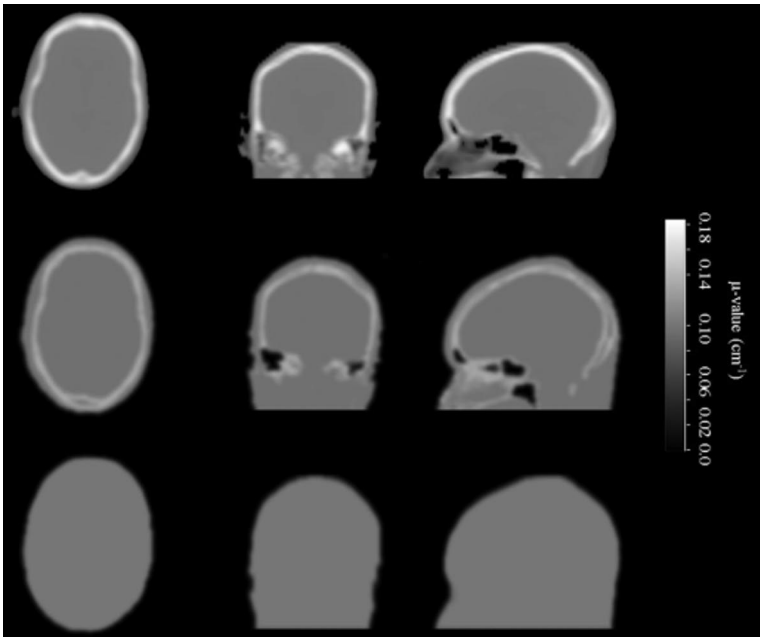

Figure 6. Example of attenuation maps. (Teuho et al., 2016), initial pipeline setup. From top to bottom, CTAC, TPB-AC, and clinical MRAC. © 2016 IEEE.

A visual evaluation of bias images from PET data reconstructed with attenuation maps of varying bone attenuation coefficients showed differences between reconstructions (Figure 7). The attenuation coefficient of 0.145 cm^{-1} produced the most balanced error distribution. The lower attenuation coefficient was underestimating the activity in most areas, and the higher coefficients were overestimating the activity across the brain. In particular, the 0.171 cm^{-1} attenuation coefficient overestimated the activity almost everywhere. The same pattern was seen in the actual PET images, where clinical attenuation map had the lowest activity, due to no increased attenuation for bone. The attenuation map reconstructed with a 0.135 cm^{-1} bone attenuation coefficient had then higher activity than the clinical attenuation map, while the CTAC had the highest activity.

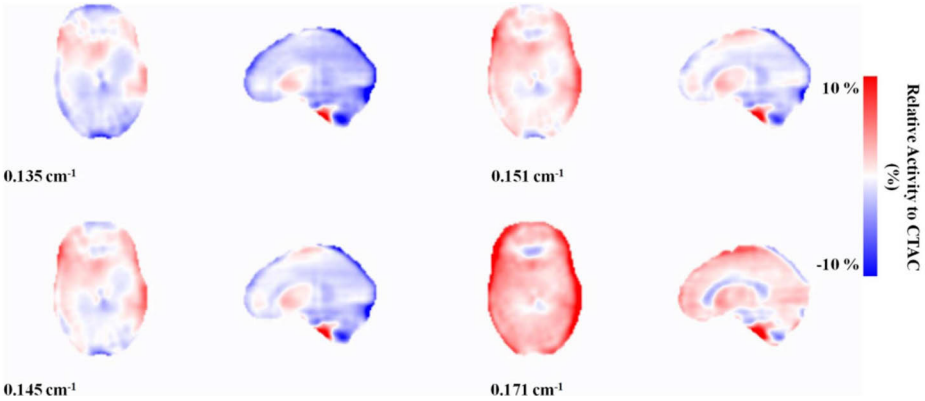

Figure 7. Example of Bias when different attenuation coefficients are used. (Teuho et al., 2016), initial pipeline setup. Bias is calculated for different bone density TPB-AC reconstructed PET images against CTAC reconstructed PET images. © 2016 IEEE.

Subject level VOI analysis kept the same order, and clinical attenuation maps yielded most PET activity underestimation, while the 0.135 cm^{-1} and 0.145 cm^{-1} attenuation maps were quite close to each other and CTAC. When the VOIs were averaged across the subjects and evaluated on VOI level, a similar pattern was detected. The clinical attenuation maps caused most underestimation, and the 0.135 cm^{-1} and 0.145 cm^{-1} attenuation maps were quite close together. With the clinical attenuation maps, 10 out of 28 regions showed at least 10% underestimation on average across the subjects. With the 0.135 cm^{-1} and 0.145 cm^{-1} attenuation maps, no region was on average more than 4% off. The 0.145 cm^{-1} attenuation coefficient map had slightly better overall mean error than the 0.135 cm^{-1} attenuation coefficient map.

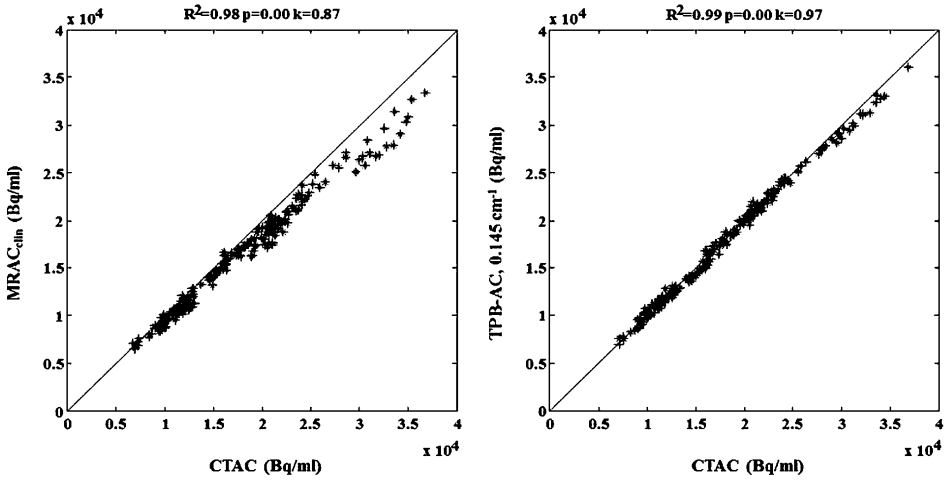

Figure 8. Scatter plots for the fit. Clinical (Left) and 0.145 cm^{-1} (right) models. The 0.135 cm^{-1} attenuation map is not shown but is closely similar to the 0.145 cm^{-1} attenuation map. (Teuho et al., 2016), initial pipeline setup © 2016 IEEE.

When looking at the scatter plots in Figure 8, the plot created from the clinical attenuation map produced visible underestimation, especially in the high activity voxels. Of the attenuation maps with separate bone tissue, this effect was not seen to the same extent. There were no clear differences in the scatter plots when different attenuation coefficients were tried for the bone tissue.

In summary, the study presented a clinically viable pipeline for MR-based attenuation correction. Different attenuation coefficients were tried, and the most suitable was selected for future studies. The created attenuation maps were computationally reasonable, and produced no systematic bias across the whole head, even though some local biases remained.

4.2 Findings – TOF effect on reconstruction quality using different AC maps

The results showed that using TOF improves the accuracy of the reconstruction in most cases. The effect was smaller when the attenuation map was more accurate. Subject level analysis showed that for all attenuation maps at least seven out of ten subjects showed reduced bias when TOF reconstruction was applied.

When all subjects were analysed together, the three class MRAC showed no clear improvement, while the two class MRACs both showed clear improvement (Figure 9). The linear regression fitting parameters are shown in Table 1. There, the R^2 values show good improvement for the two class MRAC methods. The Bland-Altman plots showed similar results, with decreased error when TOF was applied, and a generally larger effect with less accurate attenuation maps.

Table 1. Selected fitting coefficients for linear regression and R^2 values from the correlation analysis (Lindén, Teuho, Klén, et al., 2022), TOF effect evaluation.

Method		*a*	*b*	b *p*-value	*a* 95% CI	*b* 95% CI	R^2
3-class MRAC	TOF	0.966	89.1	0.142	[0.9595, 0.9717]	[−30.06, 208.32]	0.996
	non-TOF	0.960	134	0.030	[0.9544, 0.9656]	[12.77, 254.80]	0.997
2-class MRAC	TOF	0.938	152	0.115	[0.9286, 0.9479]	[−37.25, 341.09]	0.991
	non-TOF	0.932	256	0.022	[0.9216, 0.9418]	[37.25, 474.32]	0.990
2-class MRAC (no air cavities)	TOF	0.943	211	0.027	[0.9336, 0.9527]	[23.59, 398.48]	0.991
	non-TOF	0.934	320	0.003	[0.9241, 0.9439]	[104.95, 534.89]	0.990

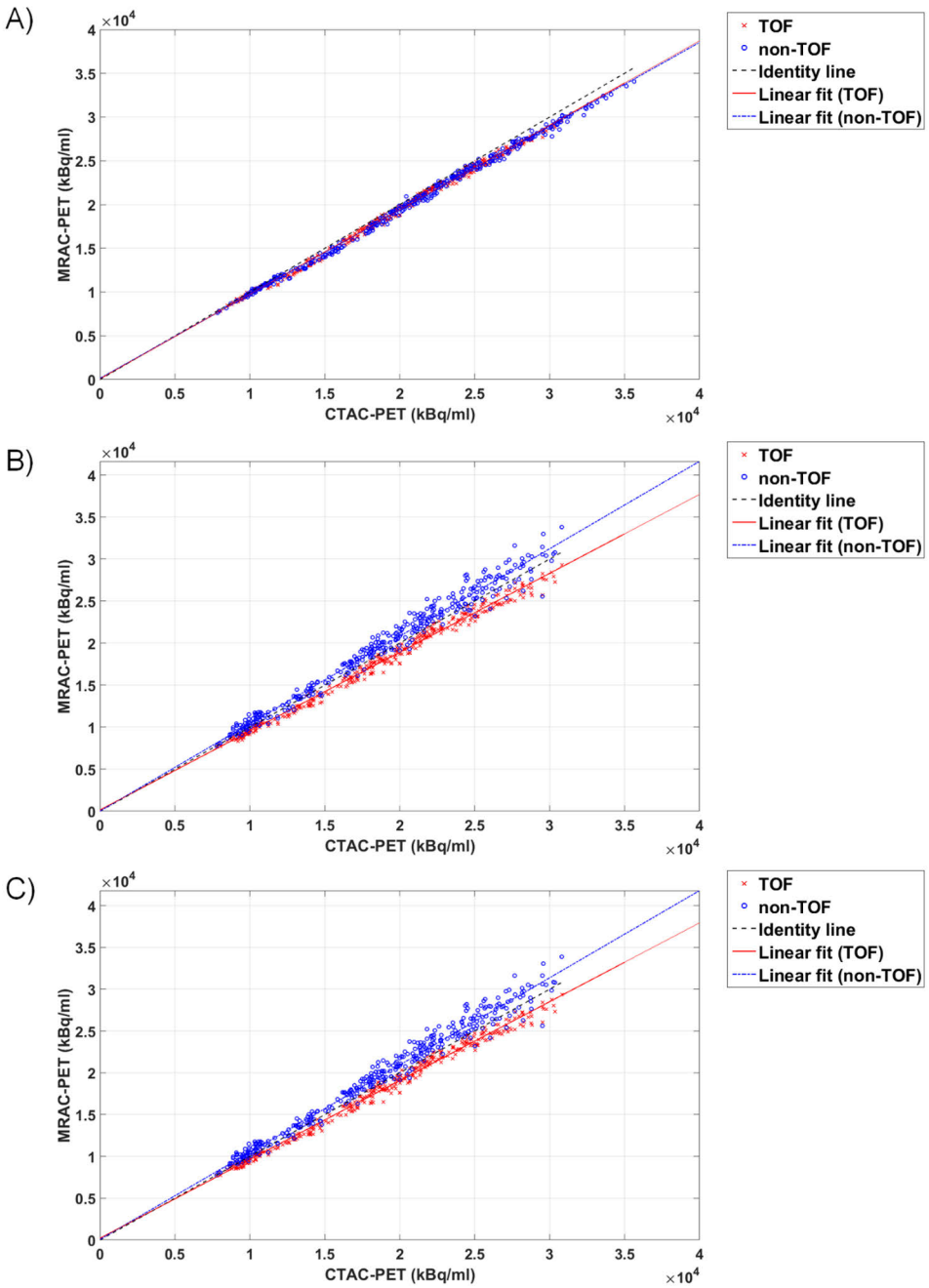

Figure 9. Regression plots of non-TOF (marked as blue 'o' symbol) and TOF (marked as red 'x' symbol) reconstructed activity values with different MRAC methods versus CTAC reconstructed PET. **A)** 3-class MRAC, **B)** 2-class MRAC, **C)** 2-class MRAC without air cavities (Lindén, Teuho, Klén, et al., 2022), TOF effect evaluation.

The mean PET images showed some local improvements when TOF reconstruction was applied. Due to small absolute changes in activity, those improvements were easier to observe with mean bias atlas images, presented in Figure 10. The bias atlas images showed improvements near sinuses where overestimation was reduced or turned into underestimation. This could be seen regardless of the attenuation map method. The two class attenuation maps had underestimation near the skull, which was to be expected when no bone class was in the model. At the same time, the middle of the brain was largely overestimated. Applying TOF reconstruction clearly helped to alleviate the biases for both the cortical and middle parts of the brain. The three-class method had more a subtle effect with slight reduction of underestimation outside of the sinus areas. In the bias standard deviation images, the cortical areas showed largest decrease of deviation for the two class methods, while the three-class method had more even reduction of deviation across the brain.

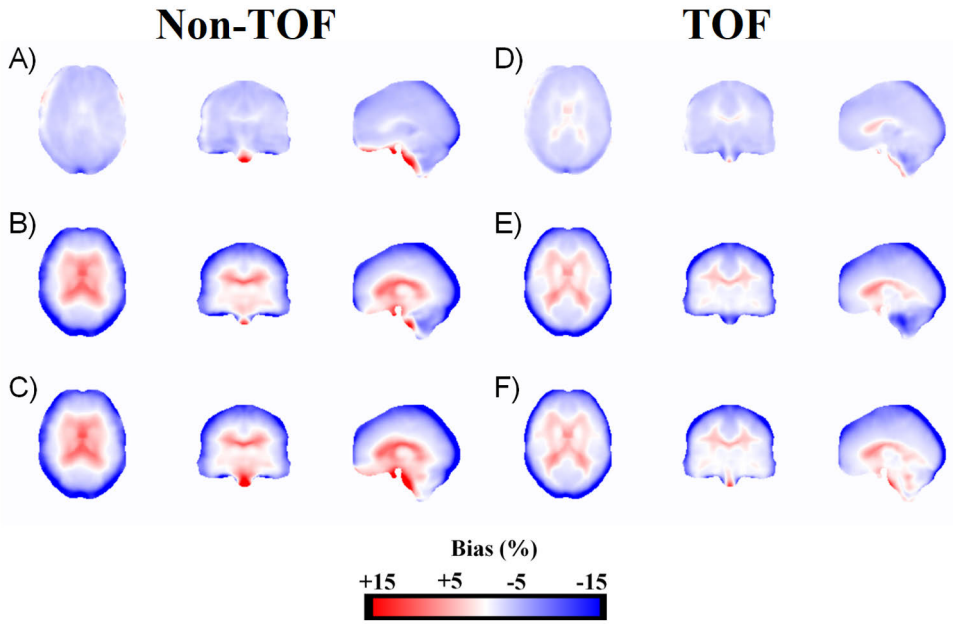

Figure 10. Mean bias atlas images from non-TOF and TOF reconstructed PET. A clear reduction of bias is seen in the TOF reconstructed images across the brain compared to non-TOF reconstructed PET. A/D) 3-class MRAC, B/E) 2-class MRAC, C/F) 2-class MRAC without air cavities using non-TOF (Left) and TOF (Right) reconstruction. (Lindén, Teuho, Klén, et al., 2022), TOF effect evaluation.

Finally, the VOI level results showed that almost two thirds of VOIs have their average bias reduced when TOF is applied with all of the two class attenuation maps.

The three-class attenuation map had a few more VOIs showing reduced bias. When the significance of the change of the median across all VOIs was tested, two out of three improvements were considered statistically significant (Table 2).

Table 2. Results of the Wilcoxon's signed rank test. Median difference denotes the median of average VOI errors in TOF vs. non-TOF reconstructions. The significance thresholds of the Benjamini–Hochberg method are given, and statistically significant p-values are denoted with *. (Lindén, Teuho, Klén, et al., 2022), TOF effect evaluation.

Method	Median Difference (TOF vs. Non-TOF)	*p*-Value (Two-Tailed)	Benjamini–Hochberg Significance Threshold
3-class MRAC	−2.9% vs. −3.3%	<0.001 *	0.0166
2-class MRAC	−5.7% vs. −5.9%	0.224	0.0500
2-class MRAC (no air cavities)	−4.7% vs. −5.4%	0.018 *	0.0333

In summary, applying TOF reconstruction improved the accuracy of the resulting PET images. The improvement was larger when the attenuation maps were less accurate. Improvement was seen in the majority of subjects and regions. However, some regions showed increased bias when TOF was applied.

4.3 Findings – Sinus segmentation method development

The results were overall favourable towards the proposed sinus methods. Visually, the air pathways and regions with mixed air and soft tissue were better delineated with both the cuboid and template methods. Using the earlier bulk density method, leaves a large amount of incorrectly assigned bone in the region. This can lead to overestimated attenuation. The mean and standard deviation PET atlas images did not show any large changes between the two new methods and old bulk density method. Some slight changes could be seen visually in occipital and parietal lobes, as well as in the cerebellum.

The DICE results of the slices containing the sinuses are presented in the Table 3. The DICE values grow over 2.5% on average, so the fit between MRAC and CTAC attenuation maps improves slightly. Of note, is that not a single subject got a worse DICE value with the new sinus methods. This indicates that the presented sinus methods remove more incorrectly assigned bone than actual bone when compared to the reference bulk density method. The correlation analysis supports the increase in DICE values, as both new sinus methods get higher bone tissue correlation factors than the bulk reference method. The DICE values alone are not sufficient, as removing actual bone can lead to too low of an attenuation correction in the region. However, the leave-one-out attenuation coefficient analysis showed that both sinus methods, together with the three-step conversion method, reduces the error of the sinus region attenuation coefficient (Figure 11).

Table 3. DICE coefficients between different methods for slices 50-70 with percentage difference to the reference. (Lindén, Teuho, Teräs, et al., 2022), sinus method evaluation.

Subject Number	Bulk	Cuboid	Template	Difference Cuboid–-Bulk	Difference Template–-Bulk
1	0.845	0.852	0.856	0.88%	1.37%
2	0.826	0.853	0.853	3.26%	3.24%
3	0.843	0.872	0.872	3.43%	3.49%
4	0.725	0.771	0.768	6.32%	5.99%
5	0.820	0.856	0.846	4.47%	3.26%
6	0.737	0.748	0.751	1.49%	1.79%
7	0.751	0.771	0.766	2.74%	2.05%
8	0.868	0.882	0.882	1.59%	1.57%
9	0.823	0.859	0.853	4.31%	3.56%
10	0.825	0.832	0.829	0.92%	0.52%
Median	0.824	0.853	0.849	3.00%	2.65%

The template method reduces the error in all ten subjects. The sinus cuboid method reduces the error in all nine subjects out of ten, but overall, the error reduction is larger than with the template method. The raw attenuation coefficient data showed overestimation for each subject with the bulk density method, whereas the sinus cuboid and template method yielded lower attenuation coefficients. While they are still in general higher than the ones in CTAC attenuation maps, a couple subjects get slightly lower attenuation coefficient with the new sinus methods than with the CTAC.

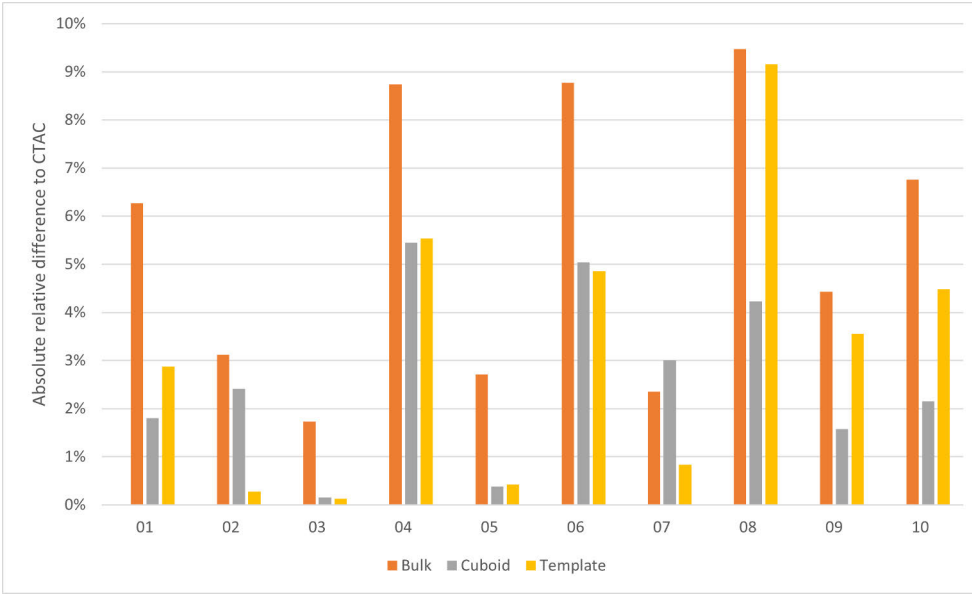

Figure 11. Absolute relative difference of the attenuation coefficients between CTAC and MRAC in the sinus region per subject. The values for the cuboid and template methods are calculated using leave-one-out method. Smaller numbers indicate a better match between MRAC and CTAC. (Lindén, Teuho, Teräs, et al., 2022), sinus method evaluation.

The PET image VOI analysis showed moderately diverging results. When TOF was not used, the bias for half of the VOIs improved, while the bias for the other half worsened, when compared to the Bulk density model. However, the absolute changes were small. When TOF was applied, 80% of the VOIs showed improvement. Thus, in TOF it can be said that the presented sinus methods reduce the bias across various brain regions on average. It is notable that for both TOF and non-TOF the cerebellum area became underestimated. The reason for the effect seen in the cerebellum is partly caused by the change of bias seen in the bias atlas images (Figure 12). The cerebellum area has spots of large overestimation along with some underestimation in the bulk density images. The presented sinus methods removed the overestimation, bringing the cerebellum to a more uniform underestimation. The sinus methods had an effect also outside the cerebellum. For both non-TOF and TOF reconstructions there was a reduction of error in the middle parts of the brain and near the sinuses.

In summary, the presented sinus methods yielded attenuation maps that have better resemblance to those created from CT images than the attenuation maps created with the previous bulk density method had achieved. This can be seen both with bone tissue delineation and the attenuation coefficient in the whole sinus area. The resulting PET images do not show conclusive VOI level changes with non-TOF reconstruction, but the vast majority of VOIs improved with TOF reconstruction. In the bias images, reduction of error can be seen in many areas, while the cerebellum remains problematic, with systemic but more uniform underestimation.

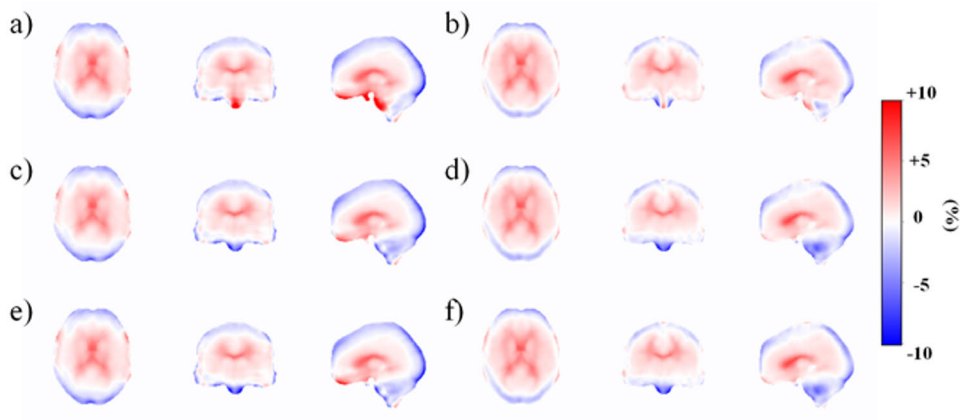

Figure 12. Bias atlas images showing the voxel-by-voxel relative difference between MRAC and CTAC based PET images in non-TOF and TOF reconstruction. Non-TOF reconstructions are represented with **a)** CTAC, **c)** bulk, **e)** cuboid and **g)** template method. TOF reconstructions are represented with **b)** CTAC, **d)** bulk, **f)** cuboid and **h)** template method. Local overestimations are seen with the bulk method whereas overestimations are reduced with the proposed methods and with TOF. (Lindén, Teuho, Teräs, et al., 2022), sinus method evaluation.

4.4 Findings – Machine learning assisted bone segmentation and HU assignment

The focus of the ML segmentation evaluation was to show that the RF method can yield better segmentation than the original SPM bone probability-based segmentation. This can be seen in the sCT comparison (Figure 13), where the continuous HU assignment RF method shows a better match with the CT.

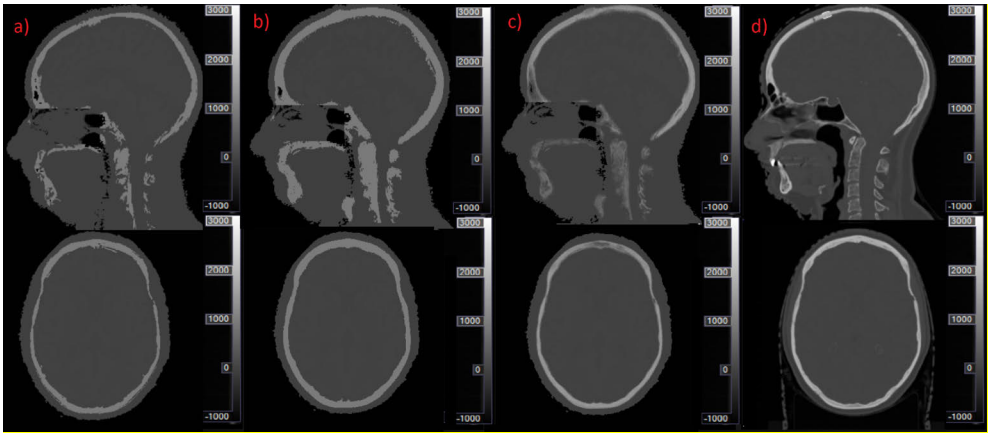

Figure 13. Comparison of the sCT images, **a)** original SPM12 based bulk density assignment sCT image. **b)** RF based bulk density assigned sCT image **c)** RF based continuous HU value assignment sCT **d)** CT. The RF sCT images are better at creating the spine and neck area, and RF continuous is also good at converting the skull bones.

The numerical results indicate that the random forest classifier is more accurate at separating bone tissue from non-bone tissue than the SPM New Segment bone probability thresholding method. The SPM Bone probability thresholding created more Type I and Type II errors than the RF based methods. This led to the random forest classifier having larger F-scores / DICE coefficients than the SPM bone probability thresholding (Table 4). Increasing the bone probability threshold would further reduce the F-scores of the bone probability thresholding model.

Table 4. Assignment score table for bone segment probability at 25% threshold (SPM Model) and for the continuous bone model together with Random Forest classifier at 27.5% threshold (RF Model). The table shows True and False rates, as well as the F-score/DICE coefficient for the classifier.

SPM Model	**Actual**		**RF Model**	**Actual**	
Prediction	**<157 HU**	**>157 HU**	**Prediction**	**<157 HU**	**>157 HU**
<157 HU	80.79 %	6.34 %	**<157 HU**	78.81 %	2.60 %
>157 HU	3.76 %	9.10 %	**>157 HU**	5.74 %	12.85 %
F-score DICE	64.3 %		**F-score DICE**	75.5 %	

The RF based method had similar MAE values with bulk assignment than the SPM method, but when the continuous bone was implemented, the RF method became the superior method (Table 5).

Table 5. MAE values for entire image / bone tissue per subject and method. RF Continous has the smallest MAE in all comparisons.

Subject	SPM12	RF Bulk	RF Continuous
1	169 / 646	182 / 644	144 / 446
2	177 / 394	196 / 426	151 / 274
3	168 / 555	184 / 572	146 / 401
4	181 / 535	198 / 559	167 / 416
5	189 / 472	190 / 454	173 / 364
6	189 / 631	187 / 554	166 / 455
7	241 / 705	232 / 597	212 / 518
8	186 / 619	192 / 621	153 / 444
9	154 / 619	183 / 617	136 / 404
10	185 / 678	199 / 649	160 / 484
Average	184 / 585	194 / 569	161 / 421

The ROC further showed the superiority of the RF model (Figure 14). The sensitivity of the RF model was constantly superior to the older bone probability thresholding method. The AUCs were 72% for SPM bone probability thresholding, and 95% for the RF classifier. Youden's J is also better for RF segmentation (0.79) than for SPM12 segmentation (0.68).

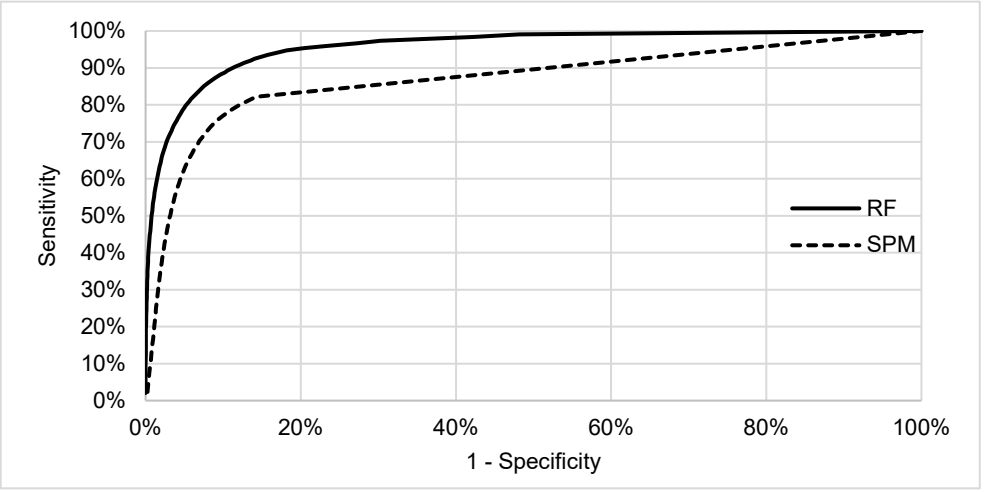

Figure 14. Receiver operating characteristics curves for RF classifier and SPM bone probability classifier. The reference SPM Bone segment probability threshold has worse performance than the new method.

In summary, the usage of machine learning techniques at the voxel level with the immediate environment as input, allowed a clear increase in the quality of the bone tissue segmentation, as well as HU assignment over the reference method. The machine learning methods had better F-scores, ROC estimates, and had smaller overall error in average HU on dataset level.

4.5 Findings – Summary

In the first part of the study, a clinically viable pipeline for MR-based attenuation correction was presented and tested. The pipeline was made modular enough to allow several of its components to be developed further. The performance changes brought by the usage of TOF were evaluated, and while overall TOF improves the performance of the reconstruction, it is not seen in all regions. When the two sinus models were presented, they both improved several aspects of the attenuation correction in the sinus area. The sinus cuboid method was overall the better performing model and performed well, especially together with TOF reconstruction. The cerebellum remained a difficult area with visible underestimation. During the ML based segmentation evaluation, the focus was shifted towards the skull and neck bones and replacing the bulk density assignment with a more dynamic model. The segmentation of bone was addressed with the use of machine learning techniques in the voxel level. The methods managed to improve the segmentation quality over the previous bone probability thresholding and provided reasonable HU values for the segmented bone tissue.

5 Discussion

This chapter discusses the findings of each part of the study separately while presenting the relevant implications between different parts of the study raised by the results.

5.1 Initial pipeline framework development

Initially, a tissue probability-based attenuation correction pipelines was successfully developed and validated for brain PET-MR. The pipeline facilitated a computationally fast and modular starting point for further improvements of attenuation correction in brain PET-MRI and MR-RT. The pipeline offered better MRAC quality than the system default, and decent accuracy when compared to similar methods in the literature at the time of the study (Navalpakkam *et al.*, 2013; Izquierdo-Garcia *et al.*, 2014; Anazodo *et al.*, 2015). Since then, some improvements have been made (Sousa *et al.*, 2020; Hamdi *et al.*, 2023).

The default settings of the New Segment functionality in SPM8 were used, and those were sufficient for this small subject group. In the pipeline, bone threshold value of 0.25 was used, which was much lower than 0.8 used in a previous study (Anazodo *et al.*, 2015). The higher value was too high in the dataset used, but there might be some system specific or post processing related reasons that affect the optimal value.

The attenuation maps generated by the pipeline match the general anatomy of the whole head and soft tissues. There are some differences in the mastoid and sinus regions, which have been problematic for MRI-based methods (Navalpakkam *et al.*, 2013). Maximum likelihood reconstruction of attenuation and activity methods could have some advantages over the methods used in these regions (Mehranian and Zaidi, 2015a, 2015c). Choosing the attenuation coefficient for bone is a trade-off between the average bias and variance. In an earlier study (Catana *et al.*, 2010), it was found that a high attenuation coefficient ($0.171 cm^{-1}$) gave the lowest bias, but had the highest variance. Meanwhile, a lower attenuation coefficient ($0.135 cm^{-1}$) was the opposite. Thus, an attenuation coefficient in the middle should be a good trade-off, and the attenuation coefficient deemed best ($0.145 cm^{-1}$) falls into this trade-off range.

Adding the bone into the attenuation map reduced the error in reconstruction. With the clinical attenuation maps, the activity levels of the whole head were underestimated by 3% to 10% with 8% as a median. When the bone was added, the errors ranged from -5% to +4%. The medians were 0% and 1% for 0.135cm^{-1} and 0.145cm^{-1} bone attenuation coefficients, respectively. On the VOI level, the clinical attenuation map had strong underestimation in most VOIs, and this was reduced clearly when the bone tissue was added to the attenuation maps. It was speculated that the usage of TOF would further reduce the error, and this was addressed by looking at the TOF effect specifically. One outlier subject went from -3% underestimation to +4% overestimation when the clinical attenuation map was changed to a one with bone. This was considered to be caused by difficulties in creating the attenuation map in sinus regions, which was addressed later in the study.

During the time of writing, the presented pipeline was rather unique, as it did not require additional CT scans or specialized MRI sequences, such as UTE or ZTE, nor did it use CR or MRI atlases for more than segmenting the images. The pipeline would not require TOF data to complete the attenuation maps either. At the same time the pipeline produced attenuation maps that were of comparable quality with previous methods (Navalpakkam *et al.*, 2013; Dickson, O'Meara and Barnes, 2014; Izquierdo-Garcia *et al.*, 2014; Anazodo *et al.*, 2015). The pipeline was aimed to be quick to compute and modular enough so further developments can easily be added. As such, the initial pipeline created a good starting point with an already very capable model.

5.2 TOF effect on reconstruction quality using different AC maps

It was shown that the use of TOF reconstruction reduces the visual and quantitative errors in MRAC in the head region. At times, though, the effects were quite small. The effect was seen across most of the head region and with all three tested attenuation maps. Of those three, two improvements were statistically significant. The usage of TOF reduced both over and underestimation compared to the reconstructions without TOF information.

The improvement was not universal, however. In approximately a third of the investigated VOIs, the absolute bias increased when TOF was used. However, the improvements were generally larger than the increases of error. At the same time, the use of TOF increased the standard deviation of the error in practically all of the brain region.

When the reconstructed activity values were compared between a CTAC reconstruction and the TOF and non-TOF reconstructions, the TOF reconstructions typically had slightly better performance. The TOF reconstruction had slightly lower

variance of activity. Since the analysis is about the match between the MRAC and CTAC reconstructions, an ideal regression curve $y = ax + b$ would have a slope coefficient a of 1 and intersect coefficient b of 0. The TOF reconstruction slightly increased the slope coefficients towards 1 and decreased the intersect parameters towards zero. Two out of three intersect parameters b changed from statistically significant to non-significant when TOF reconstruction was introduced. Although the changes were small, the TOF reconstructions slightly improved the activity modelling of the PET images. The Bland-Altman graphs showed the same tendency of error reduction with TOF reconstructions. The non-TOF data has a wider error distributions and larger average bias.

The TOF reconstruction has lower maximum uptake than the non-TOF reconstruction. This might be because of the TOF suppressing image noise more, which reduces the prevalence of extreme values in the images. The TOF reduces the local over and underestimations, and the TOF images are more aligned with the reference CTAC images, even if the effect is difficult to detect visually due to the small magnitude of the changes. The assessment of bias images confirms the pattern, as the bias can be seen decreasing across the head. Both over and underestimations are reduced in magnitude. The TOF reconstruction helped with the overestimation near the sinus regions, that was also spotted with one outlier subject when the pipeline was first developed. The effect near the sinuses was further investigated later in the study. The findings are consistent with some earlier whole-body studies (Mehranian and Zaidi, 2015b; ter Voert *et al.*, 2017) and a head study (Khalifé *et al.*, 2017). The Khalifé study used a scanner that was capable to sub 400ps timing resolution (Grant *et al.*, 2016), whereas the scanner in this study could do 525ps, and this indicates that the TOF effect can be seen in various timing resolutions and systems, not just in one. Commonly, a better timing resolution has been seen to improve overall image quality (Conti and Bendriem, 2019).

Although not every region gains improvements, on average, the results improved when TOF is used. While there is some regional variation, and to minimize the error potential, a good attenuation map remains a prerequisite for an accurate PET scan, even if the TOF information can improve the results. This was recommended by an earlier study (Emond *et al.*, 2020).

5.3 Sinus segmentation method development

Two new methods were introduced as a proof-of-concept for improved accuracy in the delineation and modelling of the attenuation coefficients in the sinus region. The effects were evaluated with a PET-MRI system, with and without TOF information. The reference method was the attenuation maps created by the original pipeline.

All subjects saw gains in accuracy when the new methods were used over the reference method. The improvement was seen in a majority of the brain region, and this became even more evident when TOF information was added. Although the model did not assign any voxels onto the bone tissue, the bone delineation was improved due to the removal of some false positives. The attenuation coefficients in the sinus region, as well as the quantitative accuracy in PET were improved.

The two methods were different from each other, and both had their own advantages and disadvantages. The cuboid method took more processing power, as an optimal cuboid is fitted to each subject's individual anatomy. In the published version of the method, computation time was saved by eliminating all tilts, so if the MRI image is sufficiently deviated from the assumption of an ideal head position, the fitting might not result in a good cuboid. However, as each subject's individual anatomy is considered, any anatomical deviations that might be difficult to handle in normalized space, are better accounted for.

The template method was computationally easy, as it only used the deformation fields that were generated during the segmentation process by default. Tilts and positional shifts were considered due to the normalization process. But as the individual anatomy was not considered in detail, anatomically deviant subjects, or those with poor template registration, might get a suboptimal sinus placement.

Both the cuboid and template method increased the DICE values in the nasal sinus region by 3% and had higher head level correlations with CTAC when compared to the reference method. Although the effect was small, every subject saw improvement. Effect can also be expected to show differences concentrated in the sinus region, as it is only a small part of the image, and the other regions are not changing from the reference method. This is supported by the almost similar whole head DICE coefficient values. There were some improvements by changing the SPM version from 8 to 12, as the reference model improved its DICE coefficients in the whole head skull tissue from 0.76±0.05 in the original to 0.85±0.04 during the sinus model evaluation.

In the attenuation coefficient analysis, the reference model performed worst, with an average error of 5.4% when cuboid method and template method yielded 2.6% and 3.2% error, respectively. While both methods were superior to the reference method, the sinus cuboid method performed better, since it potentially accounted for the individual anatomy of the subject better. The connection between the MRI intensity and CT HU values was possible to determine for the sequence used in this study. The data for any different sequence should be evaluated carefully, as the connection might not be derivable for all MRI sequences. There is a large amount of air – soft tissue borders in the sinus region which might not be adequately separable for all MRI sequences and might contain susceptibility artefacts.

When the PET reconstruction results were assessed on the VOI level, the error levels were generally small. The error levels for reference model were slightly better than when the pipeline was originally developed, likely due to the change to SPM12. The error level was close to those reported in the literature (Ouyang *et al.*, 2013; Teuho *et al.*, 2020). The best performing regions were the olfactory bulb, Heschl's gyri, lingual gyrus, and cerebellar vermis, and these regions especially benefited from a separate tissue class for the nasal sinus region. When looking at the bias atlas images, the areas near the sinuses were clearly overestimated in the reference model, and this overestimation was reduced with the new sinus modelling methods.

There were some issues in cerebellum when the cuboid and template methods were used for attenuation correction. The underestimation that was present in the VOI analysis for the reference method increased with the new methods. In bias atlas images, the reference method had overestimation in the frontal parts of the cerebellum and underestimation in the back, similarly to the earlier bias atlas images. When the new sinus methods were implemented, the frontal overestimation disappears, and the whole region was underestimated. Potentially, the thick bone near the cerebellum (Ladefoged CN *et al.*, 2015) caused problems for the attenuation maps and that lead to general underestimation of the cerebellum. There was also underestimation across the head near the skull, and this might have been caused by the bulk density assignment. In another study (Yang *et al.*, 2017), the cerebellum area improved when the sinus region attenuation was addressed, but that study used continuous coefficients for skull bone. The skull delineation was then investigated further with ML methods where a continuous model was developed.

Overall, the improvements of sinus modelling were modest, as hypothesized in another study (Sousa *et al.*, 2020). However, as some regions gained larger improvements, further developing the sinus region attenuation based on the presented methods should be beneficial for the overall performance of the brain-PET imaging.

5.4 Machine learning assisted bone segmentation and HU assignment

Earlier, there were signs that the SPM bone probability based bulk density assignment of bone tissue HU values was not optimal. This was improved by utilizing a larger amount of subject data with the help of ML techniques. As there was a large number of measurements, in the form of TPMs, available for each subject, it was natural to try machine learning methods to assist with the modelling. Each subject had two MRI sequences, in-phase, and out-phase, available, as well as CT to be used as the target variable. When those MRI images were segmented, there were also six tissue class probability maps for each image, totalling two measurements for MRI intensity and 12 tissue probabilities per subject. Finally, the

sinus cuboid information created using the process described earlier, was added as the 43rd feature. For better generalizability, it was decided to create voxel level models, rather than whole image level models.

First the RF classifier was used to perform the segmentation of bone and non-bone tissue. The RF classifier performed better than the SPM bone probability thresholding, as it had lower false positive and false negative rates, while maintaining a slightly higher true positive rate. A balance of false positives and negatives is useful, as the total volume of bone tissue in sCT images then remains close to the actual bone tissue volume. However, the false negatives are always located at the site of actual bone tissue, creating thinner than real bones. Meanwhile the false positives can be anywhere in the head. This means too little bone tissue near the cortical areas can lead to large biases in the PET AC or RTP. The RF segmentation gave a wide range of RF classifier probabilities that corresponded to near maximum DICE coefficient (Figure 15), so the threshold probability could be chosen quite freely without reducing the DICE coefficient. This was in contrast to the SPM based segmentation, where maximum DICE coefficient values were reached in very low classifier probabilities, and the DICE coefficient values declined after that (Figure 15).

In this study, the RF classifier threshold was set to 27.5%. Lowering the threshold would make the data contain more false positives, but at the same time less false negatives would incur. The value chosen as it was between the RF probability that gives the maximum DICE value and the RF probability that gives the maximum Youden's J value. This threshold was then used to create the sCT images with bulk density assignment.

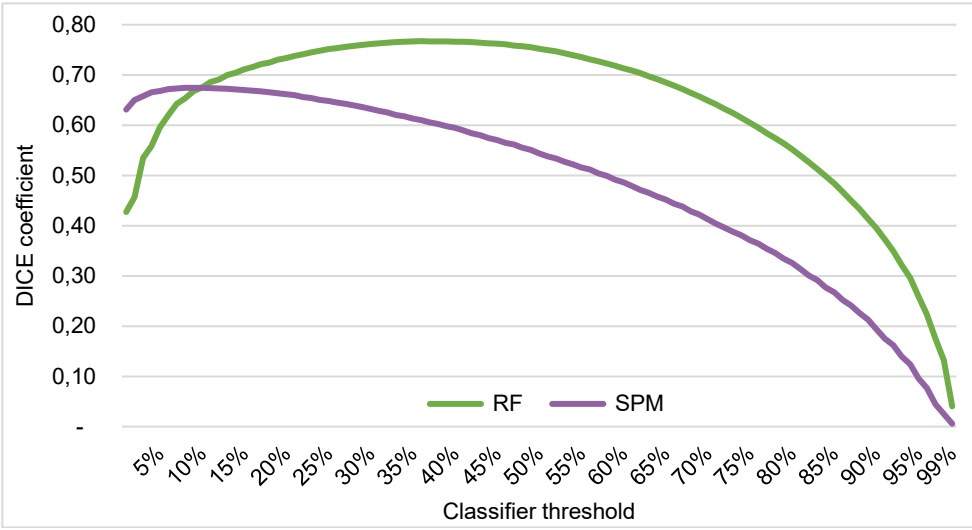

Figure 15. DICE coefficients by RF (Green) and SPM (Purple) classifier probability.

Thresholding was not an issue with the continuous HU assignment, as each voxel would get a HU value based on its classifier probability. The continuous HU values were modelling the actual HU distribution quite well. The MAE levels for the entire head and bone were clearly reduced when the RF continuous bone HU assignment was used, and the bulk density assignment method gave similar MAE values for both SPM and RF based segmentation. The whole head MAE values were comparable with those in the literature (Johansson, Karlsson and Nyholm, 2011; Gudur *et al.*, 2014; Demol *et al.*, 2016; Han, 2017), but there has as well been studies with lower MAE levels (Martinez-Girones *et al.*, 2021). Whole head MAE is always dependant on the amount of air outside the head in the compared CT and sCT images. Assuming the air outside the head gets segmented equally for CT and sCT, more air in the image, the lower the MAE is going to be. In this study, the head mask was drawn close to the skull, making the entire head MAE higher than it could be with a larger mask. The bone region MAE values between 150 and 490 have been reported by one group (Lei *et al.*, 2018; Yang *et al.*, 2019) using RF algorithms in the brain region.

The visual inspection showed typically better bone delineation for the RF methods Whereas the SPM method sCT showed bone underestimation, which is to be expected due to higher false negative rate. Due to the averaging of the RF probability to HU conversion function, the highest HU values of the CT images were not reached in the RF continuous sCT images, but the range was still substantially better than with the bulk density assignment.

Overall, the voxel-based machine learning methods were able to improve the accuracy of the bone segmentation, as well as provide another possible method to estimate the continuous HU value for the bone voxels. These updates to the segmentation should address many of the issues that arose in earlier parts of the study.

The process uses structured data with very limited sampling of the training dataset. This bypasses the typical problem with ML methods of requiring large datasets. The results were achieved with only eight subjects in the training datasets, whereas an earlier study (Ladefoged *et al.*, 2020) recommended cohort sizes of 100 to reduce the outlier effects. Some studies have used sub images in the training dataset (Andreasen *et al.*, 2015; Kläser *et al.*, 2021; Martinez-Girones *et al.*, 2021), but these sub images are still considerably larger than the sampling used in this study.

5.5 Study limitations and future outlook

Like other methods based on MRI image segmentation, the original pipeline is likely going to suffer when the subject anatomy is deformed, or there are void areas caused by metal implants in the MRI images. The ML segmentation showed reasonable performance of the bone modelling for subjects with some anatomical deviations,

but larger cohort studies with clinical data could help further assess the performance of the model in these situations. Additionally, as the pipeline should give a similar attenuation map for the same subject, even if the scans are separated by a period of time, it would also be good to assess this aspect with a test-retest study. Another limitation of the pipeline is that the processing is done externally and not on the scanner, which adds extra steps in completing the reconstruction process. One future development avenue would be to implement the pipeline into the reconstruction system, which would require support from the vendor.

It was not possible to compare the TOF and non-TOF reconstructions perfectly in this study, because of the issues reported in an earlier study (Maus *et al.*, 2015). Also, the quantitative comparison of the TOF and non-TOF reconstructions is difficult due to the differences in the reconstruction algorithms and parameters. Thus, the comparisons could only be done as TOF CTAC vs. TOF MRAC and non-TOF CTAC vs. non-TOF MRAC. Should there be a way to develop the unity between TOF and non-TOF reconstruction algorithms, the results could be revisited in a later study.

The results for TOF effect were gathered with 525ps timing resolution. There are scanners with better timing resolution (Zaidi *et al.*, 2011; Grant *et al.*, 2016; Van Sluis *et al.*, 2019; Chen *et al.*, 2021; Lee *et al.*, 2021; Schaart *et al.*, 2021), so larger improvements could be achieved by using TOF reconstructions on those scanners (Mehranian and Zaidi, 2015b). Further studies could confirm this hypothesis.

The cuboid conversion methods had some limitations. Especially as the cuboid method is reliant on the accurate segmentation of the MRI images, so that the sinus region can be located. On the other hand, the template method might struggle with changes in the individual anatomy that are not properly translated into the normalized space. It could be of interest to combine the sinus attenuation coefficient handling methods with the segmentation method similar to the ML-based approach used for the bone HU values. The process did not use RF probabilities to find the sinuses, but the SPM12 bone probabilities, like presented earlier. Another limitation is that the sinus area conversion method converts all the voxels within the sinuses to either soft tissue, air, or a mix of those two. There are some bone voxels within the sinus cuboids, they just could not be isolated with the MRI intensity-based approach that was used in the study. This was also seen with the MRI sequence used with the ML method evaluation, where the bone voxels were more prevalent than when the sinus methods were evaluated. Although the lack of bone voxels was not a large cause for error, adding a potential ML-based approach tissue probability-based handling to the sinus conversion curves could help to identify those bone voxels within the soft tissue and air mix and further improve the model.

As the sinus handling utilizes the TPMs, there are minimum requirements for the quality of the segmentation. In another study (Ranta *et al.*, 2020), the original

pipeline with some modifications was assessed with a different MRI sequence, and the segmentation worked sufficiently there. Regardless, this would need to be verified with each new sequence. Another sequence specific aspect of the sinus models was the conversion of the voxels into air and soft tissue. The conversion curve is MRI intensity based, so it needs to be re-estimated for each new MRI sequence. But after the estimation is complete, the conversion curve should work well for other subjects scanned with the same sequence, as was seen in the leave-one-out analysis of the conversion curves. Currently, the applicability of the method can be assumed to be good, but it would be interesting to try the method with some different MRI sequences.

The ML based procedure does not use complicated MRI sequences and does not make any spatial assumptions of the underlying images. This is a clear advantage, as any anatomical MRI compatible with SPM segmentation engine can be used. This should make it more adaptable to anatomically extremely different subjects than methods that rely on image or sub image level pattern recognition. The procedure should be applicable to different MRI sequences if the other sequences allow for a sufficient amount of data to be generated from the scan, and the segmentation works sufficiently well to make the quality of the data good enough for the machine learning models. The ML model was built using an RF algorithm. The model pipeline could be tested also with different algorithms, like XGBoost, or various neural networks to evaluate the accuracy of different methodology.

Current version of the pipeline works by moving the sampled data points to and from the statistical software R via comma separated values -files and requires some manual work with timing the sequential scripts on different platforms. Further work has been started to convert the pipeline into Python, allowing more streamlined pipeline. High-resolution images create large datasets, so resolution downscaling by splitting images into interleaving data files is needed when full images were converted from MRI to HU space. In this study, only the bone segment was considered. Similar studies could be performed for other tissue types, and then to create a final model that merges the different tissues together to form an sCT image, like done in the original pipeline.

These issues need to be addressed before clinical applications could be considered. When the ML model was evaluated, there were no RT plans, and the clinical performance of the improved segmentation was left for future studies. Along with a clinical evaluation, a larger cohort size should be sought, to see if there are any outlier cases where the model performance is degraded. For example, a systematic study of the effects of implants causing voids in the MRI images, could be done to assess their impact on the ML model.

Both the sinus and ML models could also be tested with a larger cohort undergoing PET imaging. Additionally, different PET-MRI devices and MRI

sequences could yield different results, and these could be tested to further assess the generalizability of the models.

Finally, the scope could be expanded to different AC methods, body parts, or removing the dependency of the SPM templates. One way to circumvent the need of an SPM segmentation could be to create a secondary model that tries to estimate the bone probability generated by the RF model using only the MRI images without the segmentation. If it was possible to do this with high enough accuracy, it would be possible to remove the SPM segmentation from the pipeline.

Currently, the training data uses MRI images segmented with SPM, but it might be of interest to use non attenuation corrected PET images. The ML model might be able to use the PET information to improve bone estimation or use that information entirely.

These advancements. could allow applications for example within PET insert MRI scans, and using the model in body parts for which SPM compatible templates are not available. It would be interesting to see if the trained multi-tissue ML model could be used as is in different body parts given the MRI sequence matches the one with which the model was trained.

Ultimately, all the improvements proposed in this study could be combined into one improved version of the original pipeline with as much automatization between steps as possible to enable clinical applicability in PET imaging and RT evaluation.

6 Conclusions

A tissue probability-based attenuation correction pipeline for PET-MRI was developed and validated with a patient study. The pipeline uses soft tissue, air, and bone tissue classes, using the New Segment functionality in SPM8. The pipeline is simple to use, modular, and quick to compute. The results indicate that the accuracy of the created attenuation correction maps was comparable to existing literature at the time of the study.

TOF reconstruction reduces the quantitative errors in the head, regardless of the quality of the MRAC. The gain from TOF reconstruction is largest when the quality of the attenuation map is low. Majority of the regions showed reduced bias with TOF information applied, but some regions show slight increase in bias. The newest scanners with time resolution around 200ps could yield larger effects than the 525ps capable one used in the TOF effect evaluation.

Two methods were introduced for improved sinus area modelling, and they provided a solid basis for further work in the sinus region attenuation correction. Both methods showed improved attenuation correction quality in the sinus region and provided an improvement to the bias near the sinus region. The usage of TOF further reduced the error in the images.

Machine learning methods were used to provide high quality separation of bone from non-bone tissues without using any specialized MRI sequences. The RF probabilities yield reasonably accurate estimates for the bone voxel HU values.

In summary, the study created a modular pipeline for MR-based attenuation correction that is usable for both MR-PET and MR-RT. In later parts of the study, a variety of improvements were studied and implemented in the pipeline. The final pipeline is easily generalizable to different MRI sequences with small additional effort. The study provides a good starting point for any future development in the area of MR-based attenuation correction in the head and neck region.

Acknowledgements

The idea to write this thesis originated during my Master's thesis work at the Turku PET Centre, and after couple years of simmering on it, I decided to take on the challenge to write it. The direction of the research morphed substantially from the first drafts to the final results, and the years I spent working on it have truly been a rewarding, albeit long, journey. I would like to take this opportunity to express my gratitude to the people who have been supporting me along the way.

First, I wish to thank my supervisors and researcher director, Professors Marko Mäkelä, Mika Teräs, Riku Klén, and Docent Jarmo Teuho for encouraging and supporting me throughout the years. Especially Riku and Jarmo, our continuous follow-up meetings have most definitely steered the project and helped me progress. Riku, I feel honoured that you have supervised all three of my theses, Bachelors, Masters, and now Doctoral. When we collaborated on my Bachelor's thesis over 12 years ago, I could not foresee this long and fruitful path together. I have also had the joy to work with Jarmo for a long time already prior to my Doctoral thesis work. You have been an instrumental in my research, along with your immense depth of knowledge, as well as, creative thinking, has greatly helped me whenever I was stuck. Mika thank you for facilitating my growth as a researcher and guiding me with your words of wisdom through the years. Additionally, I would like to thank you, Marko, for giving me inspiration and guidance early on in my studies, and later supporting me in my PhD endeavours.

Secondly, I would like to extend my gratitude to my other co-authors Jarkko Johansson, Jouni Tuisku, Terhi Tuokkola, Iiro Ranta, Jani Keyriläinen and Kaushalya Ramanayake Mudiyanselage, without whom large parts of the study would have been impossible to complete. Special thanks to Jarkko and Jouni for being there and mentoring me on my early steps towards a researcher's role. I would also like to take a moment to thank both Professors, Habib Zaidi and Jonathan Thiessen, for pre-examining my thesis. Your comments helped to get the thesis to its final form and make it better. In addition, I send my deepest gratitude to Professor, Søren Baarsgaard Hansen, for agreeing to act as my opponent.

As my thesis is somewhat between departments, I wish to thank both the Department of Mathematics and Statistics, as well as, the Turku PET Centre and

their head Professors, Iiro Honkala and Juhani Knuuti, as well as, all the staff in both departments who have provided all the help, tools and facilities needed for my research. The project was supported with financial help from The Finnish Cultural Foundation, grant from the Division of Imaging TUCH, and State Research Funding. My deepest gratitude for all the faith in me.

I want to acknowledge my friends and family for all their support. First, I want to thank Matias, Daniel, Samuli, and Markus for our friendship, and many evenings of relaxation and enjoyment between the stress of full-time work and research. Then I wish to thank my parents Jyrki and Kristiina for your continuous support and all that you have done for me over the years. Furthermore, I would also like to thank my wife's parents Paul and Lisa, who have encouraged me from across the ocean. Particularly Paul for sometimes being more excited about the project than I was.

Finally, my deepest gratitude belongs to my dearest wife, my pillar of strength, and love of my life Lauren. Without you I would not have started, let alone finished this undertaking. Words can't express how grateful I am for your love and support. Thank you for motivating and sometimes pushing me through this experience.

Turku, April 2024
Jani Lindén

List of References

Aasheim, L. B. *et al.* (2015) 'PET/MR brain imaging: evaluation of clinical UTE-based attenuation correction', *European Journal of Nuclear Medicine and Molecular Imaging*, 42(9), pp. 1439–1446. doi: 10.1007/s00259-015-3060-3.

Abella, M. *et al.* (2012) 'Accuracy of CT-based attenuation correction in PET/CT bone imaging', *Physics in Medicine & Biology*. IOP Publishing, 57(9), p. 2477. doi: 10.1088/0031-9155/57/9/2477.

Ahuja, K. *et al.* (2020) '18F-Sodium fluoride PET: History, technical feasibility, mechanism of action, normal biodistribution, and diagnostic performance in bone metastasis detection compared with other imaging modalities', *Journal of Nuclear Medicine Technology*. Society of Nuclear Medicine, pp. 9–16. doi: 10.2967/jnmt.119.234336.

Alauddin, M. M. (2012) 'Positron emission tomography (PET) imaging with 18F-based radiotracers', *American Journal of Nuclear Medicine and Molecular Imaging*. e-Century Publishing Corporation, 2(1), p. 55. Available at: /pmc/articles/PMC3478111/ (Accessed: 17 September 2023).

Anazodo, U. C. *et al.* (2015) 'Feasibility of simultaneous whole-brain imaging on an integrated PET-MRI system using an enhanced 2-point Dixon attenuation correction method', *Frontiers in neuroscience*. Front Neurosci, 8(JAN). doi: 10.3389/FNINS.2014.00434.

Andreasen, D. *et al.* (2015) 'Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain', *Medical Physics*. AAPM - American Association of Physicists in Medicine, 42(4), pp. 1596–1605. doi: 10.1118/1.4914158.

Ashburner, J. and Friston, K. J. (2005) 'Unified segmentation', *NeuroImage*. Academic Press, 26(3), pp. 839–851. doi: 10.1016/J.NEUROIMAGE.2005.02.018.

Ay, M. R. *et al.* (2011) 'Comparative assessment of energy-mapping approaches in CT-based attenuation correction for pet', *Molecular Imaging and Biology*. Springer, 13(1), pp. 187–198. doi: 10.1007/S11307-010-0303-3/METRICS.

Bai, C. *et al.* (2003) 'A Generalized Model for the Conversion from CT Numbers to Linear Attenuation Coefficients', *IEEE Transactions on Nuclear Science*, 50(5 II), pp. 1510–1515. doi: 10.1109/TNS.2003.817281.

Bailey, D. L. *et al.* (2005) *Positron Emission Tomography*. Edited by D. L. Bailey et al. London: Springer London. doi: 10.1007/B136169.

Bailey, D. L. and Willowson, K. P. (2013) 'An Evidence-Based Review of Quantitative SPECT Imaging and Potential Clinical Applications', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 54(1), pp. 83–89. doi: 10.2967/JNUMED.112.111476.

Berker, Y. and Li, Y. (2016) 'Attenuation correction in emission tomography using the emission data - A review', *Medical Physics*, 43(2), pp. 807–832. doi: 10.1118/1.4938264.

Bojorquez, J. Z. *et al.* (2017) 'What are normal relaxation times of tissues at 3 T?', *Magnetic Resonance Imaging*, pp. 69–80. doi: 10.1016/j.mri.2016.08.021.

Bowles, R. L. (1896) 'PATHOLOGICAL AND THERAPEUTIC VALUE OF THE ROENTGEN RAYS.', *The Lancet*. Elsevier, 147(3784), pp. 655–656. doi: 10.1016/S0140-6736(01)93354-9.

Breiman, L. (2001) 'Random forests', *Machine Learning*. Springer, 45(1), pp. 5–32. doi: 10.1023/A:1010933404324/METRICS.

Brooks, D. J. (2010) 'Imaging Approaches to Parkinson Disease', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 51(4), pp. 596–609. doi: 10.2967/JNUMED.108.059998.

Browne, J. and De Pierru, A. R. (1996) 'A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography', *IEEE transactions on medical imaging*. IEEE Trans Med Imaging, 15(5), pp. 687–698. doi: 10.1109/42.538946.

Burger, C. *et al.* (2002) 'PET attenuation coefficients from CT images: Experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients', *European Journal of Nuclear Medicine*. Eur J Nucl Med Mol Imaging, 29(7), pp. 922–927. doi: 10.1007/s00259-002-0796-3.

Burgos, N. *et al.* (2014) 'Attenuation correction synthesis for hybrid PET-MR scanners: Application to brain studies', *IEEE Transactions on Medical Imaging*. Institute of Electrical and Electronics Engineers Inc., 33(12), pp. 2332–2341. doi: 10.1109/TMI.2014.2340135.

Burgos, N. *et al.* (2015) 'Multi-contrast attenuation map synthesis for PET/MR scanners: assessment on FDG and Florbetapir PET tracers', *European Journal of Nuclear Medicine and Molecular Imaging*. Springer Berlin, 42(9), pp. 1447–1458. doi: 10.1007/s00259-015-3082-x.

Bushberg, J. T. *et al.* (2011) *The Essential Physics of Medical Imaging*. 3rd edn. Philadelphia: Wolters Kluwer.

Carney, J. P. J. *et al.* (2006) 'Method for transforming CT images for attenuation correction in PET/CT imaging', *Medical Physics*. John Wiley & Sons, Ltd, 33(4), pp. 976–983. doi: 10.1118/1.2174132.

Catana, C. *et al.* (2010) 'Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype', *Journal of Nuclear Medicine*, 51(9), pp. 1431–1438. doi: 10.2967/jnumed.109.069112.

Chen, K. T. *et al.* (2017) 'On the accuracy and reproducibility of a novel probabilistic atlas-based generation for calculation of head attenuation maps on integrated PET/MR scanners', *European Journal of Nuclear Medicine and Molecular Imaging*. Springer Berlin Heidelberg, 44(3), pp. 398–407. doi: 10.1007/s00259-016-3489-z.

Chen, S. *et al.* (2021) 'NEMA NU2-2012 performance measurements of the United Imaging uPMR790: an integrated PET/MR system', *European Journal of Nuclear Medicine and Molecular Imaging 2021 48:6*. Springer, 48(6), pp. 1726–1735. doi: 10.1007/S00259-020-05135-9.

Chen, X. and Liu, C. (2023) 'Deep-learning-based methods of attenuation correction for SPECT and PET', *Journal of Nuclear Cardiology*. Elsevier, 30(5), pp. 1859–1878. doi: 10.1007/S12350-022-03007-3.

Clark, J. C. *et al.* (1987) 'Current methodology for oxygen-15 production for clinical use', *International Journal of Radiation Applications and Instrumentation. Part A. Applied Radiation and Isotopes*. Pergamon, 38(8), pp. 597–600. doi: 10.1016/0883-2889(87)90122-5.

Comar, D. *et al.* (1976) 'Labelling and metabolism of methionine-methyl-11C', *European Journal of Nuclear Medicine*. Springer-Verlag, 1(1), pp. 11–14. doi: 10.1007/BF00253260/METRICS.

Confidence, R. *et al.* (2023) 'The performance of machine learning approaches for attenuation correction of PET in neuroimaging: A meta-analysis', *Journal of Neuroradiology*. Elsevier Masson, 50(3), pp. 315–326. doi: 10.1016/J.NEURAD.2023.01.157.

Conti, M. (2009) 'State of the art and challenges of time-of-flight PET', *Physica Medica: European Journal of Medical Physics*. Elsevier, 25(1), pp. 1–11. doi: 10.1016/J.EJMP.2008.10.001.

Conti, M. (2011) 'Why is TOF PET reconstruction a more robust method in the presence of inconsistent data?', *Physics in Medicine and Biology*, p. Jan 7;56(1):155-68. doi: 10.1088/0031-9155/56/1/010.

Conti, M. and Bendriem, B. (2019) 'The new opportunities for high time resolution clinical TOF PET', *Clinical and Translational Imaging*. Springer-Verlag Italia s.r.l., 7(2), pp. 139–147. doi: 10.1007/S40336-019-00316-5/FIGURES/6.

Davis, K. M. *et al.* (2020) 'PET and SPECT Imaging of the Brain: History, Technical Considerations, Applications, and Radiotracers', *Seminars in Ultrasound, CT and MRI*. W.B. Saunders, 41(6), pp. 521–529. doi: 10.1053/J.SULT.2020.08.006.

Davison, H. *et al.* (2015) 'Incorporation of Time-of-Flight Information Reduces Metal Artifacts in Simultaneous Positron Emission Tomography/Magnetic Resonance Imaging: A Simulation Study', *Investigative radiology*. Invest Radiol, 50(7), pp. 423–429. doi: 10.1097/RLI.0000000000000146.

Delso, G. *et al.* (2015) 'Clinical evaluation of zero-echo-time MR imaging for the segmentation of the skull', *Journal of Nuclear Medicine*. Society of Nuclear Medicine Inc., 56(3), pp. 417–422. doi: 10.2967/jnumed.114.149997.

Delso, G. *et al.* (2017) 'Effect of Time-of-Flight Information on PET/MR Reconstruction Artifacts: Comparison of Free-breathing versus Breath-hold MR-based Attenuation Correction', *Radiology*. Radiological Society of North America Inc., 282(1), pp. 229–235. doi: 10.1148/radiol.2016152509.

Demol, B. *et al.* (2016) 'Dosimetric characterization of MRI-only treatment planning for brain tumors in atlas-based pseudo-CT images generated from standard *T* 1-weighted MR images', *Medical Physics*. AAPM - American Association of Physicists in Medicine, 43(12), pp. 6557–6568. doi: 10.1118/1.4967480.

Dice, L. R. (1945) 'Measures of the Amount of Ecologic Association Between Species', *Ecology*. John Wiley & Sons, Ltd, 26(3), pp. 297–302. doi: 10.2307/1932409.

Dickson, J. C., O'Meara, C. and Barnes, A. (2014) 'A comparison of CT- and MR-based attenuation correction in neurological PET.', *European Journal of Nuclear Medicine and Molecular Imaging*. Springer Verlag, 41(6), pp. 1176–1189. doi: 10.1007/S00259-013-2652-Z.

Edmund, J. M. and Nyholm, T. (2017) 'A review of substitute CT generation for MRI-only radiation therapy', *Radiation Oncology*. BioMed Central Ltd., p. 28. doi: 10.1186/s13014-016-0747-y.

Emond, E. C. *et al.* (2020) 'Effect of attenuation mismatches in time of flight PET reconstruction', *Physics in medicine and biology*. Phys Med Biol, 65(8). doi: 10.1088/1361-6560/AB7A6F.

Farde, L. *et al.* (1985) 'Substituted benzamides as ligands for visualization of dopamine receptor binding in the human brain by positron emission tomography.', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 82(11), pp. 3863–3867. doi: 10.1073/PNAS.82.11.3863.

Frost, J. J. *et al.* (1985) 'Imaging Opiate Receptors in the Human Brain by Positron Tomography', *Journal of Computer Assisted Tomography*, 9(2). Available at: https://journals.lww.com/jcat/fulltext/1985/03000/imaging_opiate_receptors_in_the_human_brain_by.1.aspx.

Garnett, E. S., Firnau, G. and Nahmias, C. (1983) 'Dopamine visualized in the basal ganglia of living man', *Nature 1983 305:5930*. Nature Publishing Group, 305(5930), pp. 137–138. doi: 10.1038/305137a0.

Geurts, P., Ernst, D. and Wehenkel, L. (2006) 'Extremely randomized trees', *Machine Learning*. Springer, 63(1), pp. 3–42. doi: 10.1007/S10994-006-6226-1/METRICS.

Goitein, M. and Abrams, M. (1983) 'Multi-dimensional treatment planning: I. Delineation of anatomy', *International Journal of Radiation Oncology*Biology*Physics*. Elsevier, 9(6), pp. 777–787. doi: 10.1016/0360-3016(83)90002-0.

Goodman, M. M. *et al.* (1997) 'Fluorine-18-FPCT: a PET radiotracer for imaging dopamine transporters.', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. United States, 38(1), pp. 119–126.

Grant, A. M. *et al.* (2016) 'NEMA NU 2-2012 performance studies for the SiPM-based ToF-PET component of the GE SIGNA PET/MR system', *Medical Physics*. John Wiley & Sons, Ltd, 43(5), pp. 2334–2343. doi: 10.1118/1.4945416.

Grubbe, E. H. (1906) 'X-Rays and Radio-Active Substances as Therapeutic Agents', *Scientific American*. Springer Science and Business Media LLC, 62(1605supp), pp. 25721–25723. doi: 10.1038/SCIENTIFICAMERICAN10061906-25721BSUPP.

Grubbé, E. H. (1933) 'Priority in the Therapeutic Use of X-rays', *https://doi.org/10.1148/21.2.156*. The Radiological Society of North America , 21(2), pp. 156–162. doi: 10.1148/21.2.156.

Gudur, M. S. R. *et al.* (2014) 'A unifying probabilistic Bayesian approach to derive electron density from MRI for radiation therapy treatment planning', *Physics in Medicine and Biology*. Institute of Physics Publishing, 59(21), pp. 6595–6606. doi: 10.1088/0031-9155/59/21/6595.

Haaparanta, M. *et al.* (1996) '[18F]CFT ([18F]WIN 35,428), a radioligand to study the dopamine transporter with PET: biodistribution in rats.', *Synapse (New York, N.Y.)*. United States, 23(4), pp. 321–327. doi: 10.1002/(SICI)1098-2396(199608)23:4<321::AID-SYN10>3.0.CO;2-#.

Hamdi, M. *et al.* (2023) 'An automatic pipeline for PET/MRI attenuation correction validation in the brain', *EJNMMI Physics*. Springer Science and Business Media Deutschland GmbH, 10(1), pp. 1–14. doi: 10.1186/S40658-023-00590-3/FIGURES/5.

Han, X. (2017) 'MR-based synthetic CT generation using a deep convolutional neural network method', *Medical Physics*. John Wiley & Sons, Ltd, 44(4), pp. 1408–1419. doi: 10.1002/MP.12155.

Hitz, S. *et al.* (2014) 'Systematic comparison of the performance of integrated whole-body PET/MR imaging to conventional PET/CT for 18F-FDG brain imaging in patients examined for suspected dementia', *Journal of Nuclear Medicine*, 55(6), pp. 923–931. doi: 10.2967/jnumed.113.126813.

Hofmann, M. *et al.* (2009) 'Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques', *European journal of nuclear medicine and molecular imaging*. Eur J Nucl Med Mol Imaging, 36 Suppl 1(SUPPL. 1), pp. 93–104. doi: 10.1007/S00259-008-1007-7.

Holsti, L. R. (1995) 'Development of clinical radiotherapy since 1896', *Acta oncologica (Stockholm, Sweden)*. Acta Oncol, 34(8), pp. 995–1003. doi: 10.3109/02841869509127225.

Hounsfield, G. N. (1973) 'Computerized transverse axial scanning (tomography): I. Description of system', *British Journal of Radiology*. The British Institute of Radiology , 46(552), pp. 1016–1022. doi: 10.1259/0007-1285-46-552-1016.

Hubbell, J. (1969) 'Photon cross sections, attenuation coefficients, and energy absorption coefficients from 10keV to 100 GeV', *National Bureau of Standards, US Dept of Commerce*.

Hudson, H. M. and Larkin, R. S. (1994) 'Accelerated Image Reconstruction Using Ordered Subsets of Projection Data', *IEEE Transactions on Medical Imaging*, 13(4), pp. 601–609. doi: 10.1109/42.363108.

Ido, T. *et al.* (1978) 'Labeled 2-deoxy-D-glucose analogs. 18F-labeled 2-deoxy-2-fluoro-D-glucose, 2-deoxy-2-fluoro-D-mannose and 14C-2-deoxy-2-fluoro-D-glucose', *Journal of Labelled Compounds and Radiopharmaceuticals*. John Wiley & Sons, Ltd, 14(2), pp. 175–183. doi: 10.1002/JLCR.2580140204.

Izquierdo-Garcia, D. *et al.* (2014) 'An SPM8-Based Approach for Attenuation Correction Combining Segmentation and Nonrigid Template Formation: Application to Simultaneous PET/MR Brain Imaging', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 55(11), pp. 1825–1830. doi: 10.2967/JNUMED.113.136341.

Jack, C. R. *et al.* (2010) 'Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade', *Lancet neurology*. NIH Public Access, 9(1), p. 119. doi: 10.1016/S1474-4422(09)70299-6.

Johansson, A., Karlsson, M. and Nyholm, T. (2011) 'CT substitute derived from MRI sequences with ultrashort echo time', *Medical Physics*. John Wiley and Sons Ltd, 38(5), pp. 2708–2714. doi: 10.1118/1.3578928.

Johnstone, E. *et al.* (2018) 'Systematic Review of Synthetic Computed Tomography Generation Methodologies for Use in Magnetic Resonance Imaging–Only Radiation Therapy', *International Journal of Radiation Oncology Biology Physics*. Elsevier Inc., pp. 199–217. doi: 10.1016/j.ijrobp.2017.08.043.

Jordan, M. I. and Mitchell, T. M. (2015) 'Machine learning: Trends, perspectives, and prospects', *Science*. American Association for the Advancement of Science, 349(6245), pp. 255–260. doi: 10.1126/SCIENCE.AAA8415.

Kapanen, M. and Tenhunen, M. (2013) 'T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning', *Acta Oncologica*. Acta Oncol, 52(3), pp. 612–618. doi: 10.3109/0284186X.2012.692883.

Kemppainen, R. *et al.* (2019) 'Assessment of dosimetric and positioning accuracy of a magnetic resonance imaging-only solution for external beam radiotherapy of pelvic anatomy', *Physics and Imaging in Radiation Oncology*. Elsevier Ireland Ltd, 11, pp. 1–8. doi: 10.1016/j.phro.2019.06.001.

Khalifé, M. *et al.* (2017) 'Subject-specific bone attenuation correction for brain PET/MR: Can ZTE-MRI substitute CT scan accurately?', *Physics in Medicine and Biology*. Institute of Physics Publishing, 62(19), pp. 7814–7832. doi: 10.1088/1361-6560/aa8851.

Kinahan, P. E. *et al.* (1998) 'Attenuation correction for a combined 3D PET/CT scanner', *Medical Physics*. John Wiley & Sons, Ltd, 25(10), pp. 2046–2053. doi: 10.1118/1.598392.

Kläser, K. *et al.* (2021) 'Imitation learning for improved 3D PET/MR attenuation correction', *Medical Image Analysis*. Elsevier, 71, p. 102079. doi: 10.1016/J.MEDIA.2021.102079.

Klein, E. E. *et al.* (2009) 'Task Group 142 report: Quality assurance of medical acceleratorsa)', *Medical Physics*. John Wiley & Sons, Ltd, 36(9Part1), pp. 4197–4212. doi: 10.1118/1.3190392.

Koivula, L., Wee, L. and Korhonen, J. (2016) 'Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: Dose calculation accuracy in substitute CT images', *Medical Physics*. AAPM - American Association of Physicists in Medicine, 43(8), pp. 4634–4642. doi: 10.1118/1.4958677.

Koole, M. *et al.* (2009) 'Whole-body biodistribution and radiation dosimetry of 18F-GE067: A radioligand for in vivo brain amyloid imaging', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 50(5), pp. 818–822. doi: 10.2967/jnumed.108.060756.

Krokos, G. *et al.* (2023) 'A review of PET attenuation correction methods for PET-MR', *EJNMMI physics*. EJNMMI Phys, 10(1). doi: 10.1186/S40658-023-00569-0.

Kumar, A. and Chugani, H. T. (2013) 'The Role of Radionuclide Imaging in Epilepsy, Part 1: Sporadic Temporal and Extratemporal Lobe Epilepsy', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 54(10), pp. 1775–1781. doi: 10.2967/JNUMED.112.114397.

Ladefoged, C. N. *et al.* (2017) 'A multi-centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients', *NeuroImage*. Academic Press Inc., 147, pp. 346–359. doi: 10.1016/j.neuroimage.2016.12.010.

Ladefoged, C. N. *et al.* (2020) 'AI-driven attenuation correction for brain PET/MRI: Clinical evaluation of a dementia cohort and importance of the training group size', *NeuroImage*. Academic Press, 222, p. 117221. doi: 10.1016/J.NEUROIMAGE.2020.117221.

Ladefoged CN *et al.* (2015) 'Region specific optimization of continuous linear attenuation coefficients based on UTE (RESOLUTE): application to PET/MR brain imaging. - PubMed - NCBI', *Physics in medicine and biology*, 60(20), pp. 8047–8065. Available at: https://www.ncbi.nlm.nih.gov/pubmed/26422177 (Accessed: 8 May 2020).

Lecun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep learning', *Nature 2015 521:7553*. Nature Publishing Group, 521(7553), pp. 436–444. doi: 10.1038/nature14539.

Lederman, M. (1981) 'The early history of radiotherapy: 1895–1939', *International Journal of Radiation Oncology*Biology*Physics*. Elsevier, 7(5), pp. 639–648. doi: 10.1016/0360-3016(81)90379-5.

Lee, M. S. *et al.* (2021) 'High-resolution time-of-flight PET detector with 100 ps coincidence time resolution using a side-coupled phoswich configuration', *Physics in medicine and biology*. Phys Med Biol, 66(12). doi: 10.1088/1361-6560/AC01B5.

Lei, Y. *et al.* (2018) 'MRI-based pseudo CT synthesis using anatomical signature and alternating random forest with iterative refinement model', *Journal of Medical Imaging*. Society of Photo-Optical Instrumentation Engineers, 5(4), p. 1. doi: 10.1117/1.JMI.5.4.043504.

Lewitt, R. M. (1990) 'Multidimensional digital image representations using generalized Kaiser-Bessel window functions', *Journal of the Optical Society of America. A, Optics and image science*. J Opt Soc Am A, 7(10), p. 1834. doi: 10.1364/JOSAA.7.001834.

Lewitt, R. M. (1992) 'Alternatives to voxels for image representation in iterative reconstruction algorithms', *Physics in medicine and biology*. Phys Med Biol, 37(3), pp. 705–716. doi: 10.1088/0031-9155/37/3/015.

Li, Y. (2018) 'Optimization for Blob-Based Image Reconstruction With Generalized Kaiser–Bessel Basis Functions', *IEEE Transactions on Computational Imaging*. Institute of Electrical and Electronics Engineers (IEEE), 4(2), pp. 257–270. doi: 10.1109/TCI.2018.2796302.

Lindén, J., Teuho, J., Klén, R., *et al.* (2022) 'Are Quantitative Errors Reduced with Time-of-Flight Reconstruction When Using Imperfect MR-Based Attenuation Maps for 18F-FDG PET/MR Neuroimaging?', *Applied Sciences 2022, Vol. 12, Page 4605*. Multidisciplinary Digital Publishing Institute, 12(9), p. 4605. doi: 10.3390/APP12094605.

Lindén, J., Teuho, J., Teräs, M., *et al.* (2022) 'Evaluation of three methods for delineation and attenuation estimation of the sinus region in MR-based attenuation correction for brain PET-MR imaging', *BMC Medical Imaging*. BioMed Central Ltd, 22(1), pp. 1–14. doi: 10.1186/S12880-022-00770-0/FIGURES/7.

Love, C. and Palestro, C. J. (2004) 'Radionuclide Imaging of Infection', *Journal of Nuclear Medicine Technology*, 32(2), pp. 47 LP – 57. Available at: http://tech.snmjournals.org/content/32/2/47.abstract.

Mahar, R. *et al.* (2018) 'Quantitative measurement of T2, T1ρ and T1 relaxation times in articular cartilage and cartilage-bone interface by SE and UTE imaging at microscopic resolution', *Journal of Magnetic Resonance*. Academic Press Inc., 297, pp. 76–85. doi: 10.1016/j.jmr.2018.10.008.

Martinez-Girones, P. M. *et al.* (2021) 'Franken-CT: Head and Neck MR-Based Pseudo-CT Synthesis Using Diverse Anatomical Overlapping MR-CT Scans', *Applied Sciences 2021, Vol. 11, Page 3508*. Multidisciplinary Digital Publishing Institute, 11(8), p. 3508. doi: 10.3390/APP11083508.

Martinez-Möller, A. and Nekolla, S. G. (2012) 'Attenuation correction for PET/MR: Problems, novel approaches and practical solutions', *Zeitschrift fur Medizinische Physik*, 22(4), pp. 299–310. doi: 10.1016/j.zemedi.2012.08.003.

Matej, S. and Lewitt, R. M. (1995) 'Efficient 3D Grids for Image Reconstruction Using Spherically-Symmetric Volume Elements', *IEEE Transactions on Nuclear Science*, 42(4), pp. 1361–1370. doi: 10.1109/23.467854.

Matej, S. and Lewitt, R. M. (1996) 'Practical considerations for 3-D image reconstruction using spherically symmetric volume elements', *IEEE Transactions on Medical Imaging*, 15(1), pp. 68–78. doi: 10.1109/42.481442.

Maus, J. *et al.* (2015) 'Evaluation of in vivo quantification accuracy of the Ingenuity-TF PET/MR', *Medical Physics*, 42(10). doi: 10.1118/1.4929936.

McNaughton, J. *et al.* (2023) 'Machine Learning for Medical Image Translation: A Systematic Review', *Bioengineering 2023, Vol. 10, Page 1078*. Multidisciplinary Digital Publishing Institute, 10(9), p. 1078. doi: 10.3390/BIOENGINEERING10091078.

Mecheter, I. *et al.* (2020) 'MR Image-Based Attenuation Correction of Brain PET Imaging: Review of Literature on Machine Learning Approaches for Segmentation', *Journal of Digital Imaging*. Springer, pp. 1–18. doi: 10.1007/s10278-020-00361-x.

Mehranian, A., Arabi, H. and Zaidi, H. (2016a) 'Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI', *NeuroImage*. Academic Press Inc., 130, pp. 123–133. doi: 10.1016/j.neuroimage.2016.01.060.

Mehranian, A., Arabi, H. and Zaidi, H. (2016b) 'Vision 20/20: Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities', *Medical Physics*. AAPM - American Association of Physicists in Medicine, 43(3), pp. 1130–1155. doi: 10.1118/1.4941014.

Mehranian, A. and Zaidi, H. (2015a) 'Clinical Assessment of Emission- and Segmentation-Based MR-Guided Attenuation Correction in Whole-Body Time-of-Flight PET/MR Imaging', *Journal of*

Nuclear Medicine. Society of Nuclear Medicine, 56(6), pp. 877–883. doi: 10.2967/JNUMED.115.154807.

Mehranian, A. and Zaidi, H. (2015b) 'Impact of time-of-flight PET on quantification errors in MR imaging-based attenuation correction', *Journal of Nuclear Medicine*. Society of Nuclear Medicine Inc., 56(4), pp. 635–641. doi: 10.2967/jnumed.114.148817.

Mehranian, A. and Zaidi, H. (2015c) 'Joint Estimation of Activity and Attenuation in Whole-Body TOF PET/MRI Using Constrained Gaussian Mixture Models', *IEEE TRANSACTIONS ON MEDICAL IMAGING*, 34(9). doi: 10.1109/TMI.2015.2409157.

Mehranian, A., Zaidi, H. and Reader, A. J. (2017) 'MR-guided joint reconstruction of activity and attenuation in brain PET-MR', *NeuroImage*. Academic Press Inc., 162, pp. 276–288. doi: 10.1016/j.neuroimage.2017.09.006.

Mérida, I. *et al.* (2017) 'Multi-atlas attenuation correction supports full quantification of static and dynamic brain PET data in PET-MR', *Physics in Medicine and Biology*. Institute of Physics Publishing, 62(7), pp. 2834–2858. doi: 10.1088/1361-6560/aa5f6c.

Metcalfe, P. *et al.* (2013) 'The Potential for an Enhanced Role for MRI in Radiation-therapy Treatment Planning', *Technology in Cancer Research & Treatment*. SAGE Publications, 12(5), p. 429. doi: 10.7785/TCRT.2012.500342.

Nag, S. *et al.* (2000) 'The American Brachytherapy Society recommendations for high-dose-rate brachytherapy for carcinoma of the cervix', *International Journal of Radiation Oncology*Biology*Physics*. Elsevier, 48(1), pp. 201–211. doi: 10.1016/S0360-3016(00)00497-1.

Navalpakkam, B. K. *et al.* (2013) 'Magnetic resonance-based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps', *Investigative Radiology*. doi: 10.1097/RLI.0b013e318283292f.

Nüsslin, F. (2020) 'Wilhelm Conrad Röntgen: The scientist and his discovery', *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics (AIFB)*. Phys Med, 79, pp. 65–68. doi: 10.1016/J.EJMP.2020.10.010.

Nutt, R. (2002) 'The History of Positron Emission Tomography', *Molecular Imaging & Biology*. No longer published by Elsevier, 4(1), pp. 11–26. doi: 10.1016/S1095-0397(00)00051-0.

Nuyts, J. *et al.* (1999) 'Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms', *IEEE Transactions on Medical Imaging*. IEEE, 18(5), pp. 393–403. doi: 10.1109/42.774167.

Ouyang, J. *et al.* (2013) 'Bias atlases for segmentation-based pet attenuation correction using PET-CT and MR', *IEEE Transactions on Nuclear Science*, 60(5), pp. 3373–3382. doi: 10.1109/TNS.2013.2278624.

Owrangi, A. M., Greer, P. B. and Glide-Hurst, C. K. (2018) 'MRI-only treatment planning: Benefits and challenges', *Physics in Medicine and Biology*. Institute of Physics Publishing, p. 05TR01. doi: 10.1088/1361-6560/aaaca4.

Persson, A. *et al.* (1985) 'Imaging of [11C]-labelled RO 15-1788 binding to benzodiazepine receptors in the human brain by positron emission tomography', *Journal of Psychiatric Research*. J Psychiatr Res, 19(4), pp. 609–622. doi: 10.1016/0022-3956(85)90080-9.

Pollard, J. M. *et al.* (2017) 'The future of image-guided radiotherapy will be MR guided', *The British Journal of Radiology*. British Institute of Radiology, 90(1073). doi: 10.1259/BJR.20160667.

Purdy, J. A. (2001) 'Intensity-modulated radiotherapy: current status and issues of interest', *International Journal of Radiation Oncology*Biology*Physics*. Elsevier, 51(4), pp. 880–914. doi: 10.1016/S0360-3016(01)01749-7.

Purdy, J. A. (2008) 'Dose to normal tissues outside the radiation therapy patient's treated volume: A review of different radiation therapy techniques', *Health Physics*, 95(5), pp. 666–676. doi: 10.1097/01.HP.0000326342.47348.06.

Ranta, I. *et al.* (2020) 'Assessment of MRI-Based Attenuation Correction for MRI-Only Radiotherapy Treatment Planning of the Brain', *Diagnostics*, p. 10(5):299. doi: 10.3390/diagnostics10050299.

Reader, A. J. and Zaidi, H. (2007) 'Advances in PET Image Reconstruction', *PET Clinics*. Elsevier, 2(2), pp. 173–190. doi: 10.1016/J.CPET.2007.08.001.

Rezaei, A. *et al.* (2012) 'Simultaneous reconstruction of activity and attenuation in time-of-flight PET', *IEEE Transactions on Medical Imaging*, 31(12), pp. 2224–2233. doi: 10.1109/TMI.2012.2212719.

Röntgen, W. C. (1896) 'ON A NEW KIND OF RAYS', *Science (New York, N.Y.)*. Science, 3(59), pp. 227–231. doi: 10.1126/SCIENCE.3.59.227.

Rosenow, U. F. (1995) 'Notes on the legacy of the Röntgen rays', *Medical physics*. Med Phys, 22(11 Pt 2), pp. 1855–1867. doi: 10.1118/1.597641.

Ryzhikov, N. N. *et al.* (2005) 'Preparation of highly specific radioactivity [18F]flumazenil and its evaluation in cynomolgus monkey by positron emission tomography', *Nuclear Medicine and Biology*. Elsevier, 32(2), pp. 109–116. doi: 10.1016/j.nucmedbio.2004.11.001.

Sartor, O. *et al.* (2017) 'An exploratory analysis of alkaline phosphatase, lactate dehydrogenase, and prostate-specific antigen dynamics in the phase 3 ALSYMPCA trial with radium-223', *Annals of Oncology*. Elsevier, 28(5), pp. 1090–1097. doi: 10.1093/ANNONC/MDX044.

Schaart, D. R. *et al.* (2021) 'Time of Flight in Perspective: Instrumental and Computational Aspects of Time Resolution in Positron Emission Tomography', *IEEE transactions on radiation and plasma medical sciences*. NIH Public Access, 5(5), p. 598. doi: 10.1109/TRPMS.2021.3084539.

Schramm, G. *et al.* (2013) 'Quantitative accuracy of attenuation correction in the Philips Ingenuity TF whole-body PET/MR system: a direct comparison with transmission-based attenuation correction', *Magma (New York, N.Y.)*. MAGMA, 26(1), pp. 115–126. doi: 10.1007/S10334-012-0328-5.

Serai, S. D. (2022) 'Basics of magnetic resonance imaging and quantitative parameters T1, T2, T2*, T1rho and diffusion-weighted imaging', *Pediatric radiology*. Pediatr Radiol, 52(2), pp. 217–227. doi: 10.1007/S00247-021-05042-7.

Shao, Y. *et al.* (1997) 'Simultaneous PET and MR imaging', *Physics in medicine and biology*. Phys Med Biol, 42(10), pp. 1965–1970. doi: 10.1088/0031-9155/42/10/010.

Shepp, L. A. and Vardi, Y. (1982) 'Maximum Likelihood Reconstruction for Emission Tomography', *IEEE Transactions on Medical Imaging*, 1(2), pp. 113–122. doi: 10.1109/TMI.1982.4307558.

Van Sluis, J. *et al.* (2019) 'Performance Characteristics of the Digital Biograph Vision PET/CT System', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 60(7), pp. 1031–1036. doi: 10.2967/JNUMED.118.215418.

Son, Y. D. *et al.* (2010) 'Analysis of biased PET images caused by inaccurate attenuation coefficients.', *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. J Nucl Med, 51(5), pp. 753–760. doi: 10.2967/jnumed.109.070326.

Sousa, J. M. *et al.* (2020) 'Accuracy and precision of zero-echo-time, single- and multi-atlas attenuation correction for dynamic [11C]PE2I PET-MR brain imaging', *EJNMMI Physics*. Springer Science and Business Media Deutschland GmbH, 7(1), p. 77. doi: 10.1186/s40658-020-00347-2.

Surti, S. and Karp, J. S. (2016) 'Advances in time-of-flight PET', *Physica Medica*. Elsevier, 32(1), pp. 12–22. doi: 10.1016/J.EJMP.2015.12.007.

Svirydenka, H. *et al.* (2017) 'The Effect of Susceptibility Artifacts Related to Metallic Implants on Adjacent-Lesion Assessment in Simultaneous TOF PET/MR', *J Nucl Med*, 58(7), pp. 1167–1173. doi: 10.2967/jnumed.116.180802.

Tapia, E. (2011) 'A note on the computation of high-dimensional integral images', *Pattern Recognition Letters*. Elsevier Science Inc. PUB552 New York, NY, USA , 32(2), pp. 197–201. doi: 10.1016/j.patrec.2010.10.007.

Teuho, J. *et al.* (2016) 'Tissue Probability-Based Attenuation Correction for Brain PET/MR by Using SPM8', *IEEE Transactions on Nuclear Science*, 63(5), pp. 2452–2463. doi: 10.1109/TNS.2015.2513064.

Teuho, J. *et al.* (2017) 'Effect of Brain Tissue and Continuous Template-Based Skull in MR-Based Attenuation Correction for Brain PET/MR', *IEEE Transactions on Radiation and Plasma Medical Sciences*, 1(3), pp. 246–261. doi: 10.1109/tns.2017.2692306.

Teuho, J. *et al.* (2020) 'Magnetic Resonance-Based Attenuation Correction and Scatter Correction in Neurological Positron Emission Tomography/Magnetic Resonance Imaging—Current Status With Emerging Applications', *Frontiers in Physics*. Frontiers Media S.A., p. 243. doi: 10.3389/fphy.2019.00243.

Townsend, D. W. *et al.* (1999) 'SMART scanner: a combined PET/CT tomograph for clinical oncology', *IEEE Nuclear Science Symposium and Medical Imaging Conference*. IEEE, 2, pp. 1170–1174. doi: 10.1109/NSSMIC.1998.774368.

Valverde, S. *et al.* (2017) 'Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach', *NeuroImage*. Academic Press, 155, pp. 159–168. doi: 10.1016/J.NEUROIMAGE.2017.04.034.

Vandenberghe, S. *et al.* (2016) 'Recent developments in time-of-flight PET', *EJNMMI Physics 2016 3:1*. SpringerOpen, 3(1), pp. 1–30. doi: 10.1186/S40658-016-0138-3.

Vandenberghe, S. and Karp, J. (2006) 'Rebinning and reconstruction techniques for 3D TOF-PET', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. North-Holland, 569(2), pp. 421–424. doi: 10.1016/J.NIMA.2006.08.065.

ter Voert, E. E. G. W. *et al.* (2017) 'The Effect of Defective PET Detectors in Clinical Simultaneous [18 F]FDG Time-of-Flight PET/MR Imaging', *Molecular imaging and biology*. Mol Imaging Biol, 19(4), pp. 626–635. doi: 10.1007/S11307-016-1023-0.

Wagenknecht, G. *et al.* (2013) 'MRI for attenuation correction in PET: methods and challenges', *Magma (New York, N.Y.)*. MAGMA, 26(1), pp. 99–113. doi: 10.1007/S10334-012-0353-4.

Wilson, A. A. *et al.* (2000) 'Novel radiotracers for imaging the serotonin transporter by positron emission tomography: Synthesis, radiosynthesis, and in vitro and ex vivo evaluation of 11C-labeled 2-(phenylthio)araalkylamines', *Journal of Medicinal Chemistry*. J Med Chem, 43(16), pp. 3103–3110. doi: 10.1021/jm000079i.

Wong, W. H. *et al.* (1983) 'Image Improvement and Design Optimization of the Time-of-Flight PET', *Journal of Nuclear Medicine*. Society of Nuclear Medicine, 24(1), pp. 52–60. Available at: https://jnm.snmjournals.org/content/24/1/52 (Accessed: 7 December 2023).

Yang, J. *et al.* (2017) 'Evaluation of sinus/edge-corrected zero-echo-time–Based attenuation correction in brain PET/MRI', *Journal of Nuclear Medicine*. Society of Nuclear Medicine Inc., 58(11), pp. 1873–1879. doi: 10.2967/jnumed.116.188268.

Yang, X. *et al.* (2019) 'MRI-based attenuation correction for brain PET/MRI based on anatomic signature and machine learning', *Physics in Medicine & Biology*. IOP Publishing, 64(2), p. 025001. doi: 10.1088/1361-6560/AAF5E0.

Youden, W. J. (1950) 'Index for rating diagnostic tests.', *Cancer*. United States, 3(1), pp. 32–35. doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3.

Yu, V. Y. *et al.* (2021) 'A multi-institutional analysis of a general pelvis continuous Hounsfield unit synthetic CT software for radiotherapy', *Journal of Applied Clinical Medical Physics*. John Wiley and Sons Ltd, 22(3), pp. 207–215. doi: 10.1002/acm2.13205.

Zaidi, H. *et al.* (2011) 'Design and performance evaluation of a whole-body Ingenuity TF PET-MRI system', *Physics in Medicine and Biology*. doi: 10.1088/0031-9155/56/10/013.

Zhou, L. *et al.* (2017) 'Machine learning on big data: Opportunities and challenges', *Neurocomputing*. Elsevier, 237, pp. 350–361. doi: 10.1016/J.NEUCOM.2017.01.026.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama Oy, Turku, Finland 2024

ISBN 978-951-29-9654-4 (PRINT)
ISBN 978-951-29-9655-1 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)