Aus der
Radiologischen Universitätsklinik Tübingen
Abteilung Diagnostische und Interventionelle Radiologie
Sektion für Experimentelle Radiologie

Standardized MR-based localization and quantification of adipose tissue compartments in large cohorts of healthy individuals

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Humanwissenschaften

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von
Haueise, Tobias
2025

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Professor Dr. J. Machann
2. Berichterstatter: Privatdozentin Dr. G. Hagberg

Tag der Disputation: 09.04.2025

Für meine Eltern,
Eva und Thomas Haueise

Table of contents

Acronyms

AT	adipose tissue
BMAT	Bone marrow adipose tissue
BMI	Body mass index
CT	Computed tomography
DL	Deep learning
DZD	German Center for Diabetes Research
DXA	Dual energy X-ray absorptiometry
EPIC	European Prospective Investigation into Cancer and Nutrition
FF	Fat fraction
IP	In-phase
KORA	Cooperative Health Research in the Region of Augsburg
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
NAKO	German National Cohort
OP	Opposed-phase
PDFF	Proton density fat fraction
PLIS	Prediabetes Lifestyle Intervention Study
SAT	Subcutaneous adipose tissue
SHIP	Study of Health In Pomerania
TDFS	Tuebingen Diabetes Family Study
TULIP	Tuebingen Lifestyle Intervention Program
T2D	Type 2 diabetes
UKB	UK Biobank
VAT	Visceral adipose tissue
VIBE	Volume interpolated breath-hold examination
WHO	World Health Organization

List of Figures

List of Tables

1 Introduction

Obesity, often phenotypically described as the complication of too much adipose tissue (Cypess 2022), is a global public health challenge and a heterogeneous chronic disease being associated with disorders of, for example, the endocrine or cardiovascular system. In 2022, 2.5 billion adults were overweight, of these 890 million were living with obesity (World Health Organization 2024). Following the current international recommendation of the World Health Organization (WHO) for the diagnosis of obesity, adults are classified based on the body mass index (BMI, calculated by the ratio of weight in kilograms and squared height in meters) as overweight (BMI$\geq$25kg/m²) or obese (BMI$\geq$30kg/m²). According to the WHO guidelines, BMI is a surrogate marker and additional measurements such as the waist circumference could help the diagnosis (World Health Organization 2024), but BMI alone does not provide insights into the heterogeneity of the disease and an application to the diagnosis of (local) excess adiposity in the individual patient is limited (Bray 2023). Besides other critique, for example, not accounting for taller populations today compared to previous generations thereby underestimating the health burden of obesity (Stefan et al. 2024), resulting from its anthropometry-based definition, BMI is not a very good measure for the quantity of adipose tissue in body, and more importantly, BMI is not able to differentiate regional adipose tissue compartments and to assess their spatial distribution within the body.

However, the important role of adipose tissue distribution for the characterization of different types of obesity has already been described by Jean Vague in the 1950s concluding that a subtype of obesity with upper body predominance is associated to metabolic disorders (Vague 1956). Vague's findings were reinforced in the 1980s by Kissebah et al. who suggested that "a classification of obesity based upon the pattern of body fat distribution may provide a means of assessing the susceptibility of obese women to glucose intolerance, hyperinsulinemia, and hypertriglyceridemia" (Kissebah et al. 1982). These initial findings relating obesity of upper body predominance, referred to as central or abdominal obesity, have been confirmed by several other studies in the following years (Krotkiewski et al. 1983; Evans et al. 1984; Lapidus et al. 1984; Després et al. 1989). First prospective population-based studies have found associations of abdominal obesity and the risk of developing diabetes (Ohlson et al. 1985), "particularly in a cohort with a BMI that was not in the range of obesity at the time of baseline

examination" (Gupta et al. 2024).

All these early studies relied on anthropometric assessment of adipose tissue distribution, for example, using waist-to-hip ratio measured by tape or using calipers, but it was also in the 1980s that more precise imaging-based measures have been developed. All these approaches have in common that they allow for differentiation of adipose tissue compartments into subcutaneous adipose tissue (SAT) and internal adipose tissue. Internal adipose tissue can further be classified into non-visceral internal adipose tissue (for example, paravertebral adipose tissue) and visceral adipose tissue (VAT) (Shen et al. 2003). If VAT is defined according to the origin of the term viscera, it refers to adipose tissue distributed in the three ventral body cavities, i.e., intrathoracic, intraabdominal, and intrapelvic adipose tissue. However, VAT commonly refers to the sum of intraabdominal and intrapelvic adipose tissue which is reasonable as the pelvic and abdominal cavity are anatomically connected (see Fig. 1.1). If necessary, Shen et al. propose a more detailed subclassification into intra- and extraperitoneal components. Intrathoracic adipose tissue is mainly distributed around the heart(Shen et al. 2003) and can further be subclassified epicardial and pericardial adipose tissue (Iacobellis 2022). Detailed differentiation of adipose tissue compartments and their quantification thus allow for a more precise analysis of body composition.

In the first years, computed tomography (CT) imaging was initially used for the localization and quantification of abdominal adipose tissue (Borkan et al. 1982) but was soon extended to the whole body (Tokunaga et al. 1983). Thereafter, imaging-based adipose tissue distribution was linked to glucose tolerance and diabetes (Sparrow et al. 1986; Shuman et al. 1986). One major drawback of CT is the exposure to ionizing radiation, limiting its application in intervention studies with multiple visits, for example, effects of diet or exercise, and in cohorts of healthy subjects. The development of imaging-based measures of body composition further involved dual energy X-ray absorptiometry (DXA) (Svendsen et al. 1993; Jensen et al. 1995). Similar to CT, DXA involves exposure to ionizing radiation, and therefore, its (repeated) application in research, especially regarding healthy participants, requires legal and ethical justification as well as consideration of appropriate alternatives.

In 1989, magnetic resonance imaging (MRI) was proposed as a non-invasive and radiation-free alternative for the analysis of body composition. In a pioneering work, Staten et al. acquired three axial slices at three body locations to allow repeated measurements on a subject to determine changes in specific adipose tissue compartments during an intervention (Staten et al. 1989). Similar to the

applications of CT, MRI of the whole body was introduced to enable more precise measurements of adipose tissue distribution, particularly of VAT, and its relationship to anthropometry: in 1992, Ross et al. first acquired 10 mm-slices every 50 mm from head to feet in 27 men and 15 women (Ross et al. 1992; Ross et al. 1993). Additionally, the relationship between different adipose tissue compartments and the interindividual phenotypic variability of adipose tissue distribution in the whole body attracted increased interest (Thomas et al. 1998).

In the following years, numerous studies have implemented imaging-based body composition analyses aiming not only to differentiate SAT and VAT but also to investigate smaller adipose tissue depots and ectopic fat accumulations, i.e., the storage of lipids in tissues other than adipose tissue (Snel et al. 2012), for example in hepatic or pancreatic tissue measured as proton density fat fraction (PDFF) assessed by advanced quantitative chemical shift encoded imaging (Reeder et al. 2010; Zhong et al. 2014). Additionally, bone marrow adipose tissue (BMAT), which is considered a secretory (secreting growth- and cell differentiation regulating cytokines) and metabolic organ (Pachón-Peña et al. 2022), has been included in these analyses. BMAT is a distinct subtype of adipose tissue (Suchacki et al. 2020) influenced by and influencing metabolic factors, such as energy and bone metabolism (Pachón-Peña et al. 2022; Shu et al. 2022), and is responding to nutritional interventions, for example expansion during caloric restriction (Pachón-Peña et al. 2022).

Today, body composition analysis by MR-based localization and quantification of adipose tissue compartments, using standardized definitions and delineations of body regions and anatomical structures, can provide precise and objective measures of the heterogeneous distribution of adipose tissue. For example, Thomas et al. proposed a new thin-on-the-outside fat-on-the-inside subphenotype with increased metabolic risk (Thomas et al. 2012). In combination with metabolic examinations using, for example, oral glucose tolerance tests or euglycemic-hyperinsulinemic and hyperglycemic clamps, knowledge from adipose tissue distribution can be used for metabolic and cardiovascular risk prediction and identification of independent risk factors for type 2 diabetes and cardiovascular diseases (Neeland et al. 2019; Wagner et al. 2022) by identification of related subphenotypes of adipose tissue distribution (Yamazaki et al. 2022). Furthermore, cluster analyses have been used to identify subphenotypes of prediabetes which could be characterized, among other metabolic variables, by single adipose tissue compartments (i.e., VAT, renal sinus fat, and liver fat) and a metabolically healthy subtype of obesity (Wagner et al. 2021; Stefan 2008).

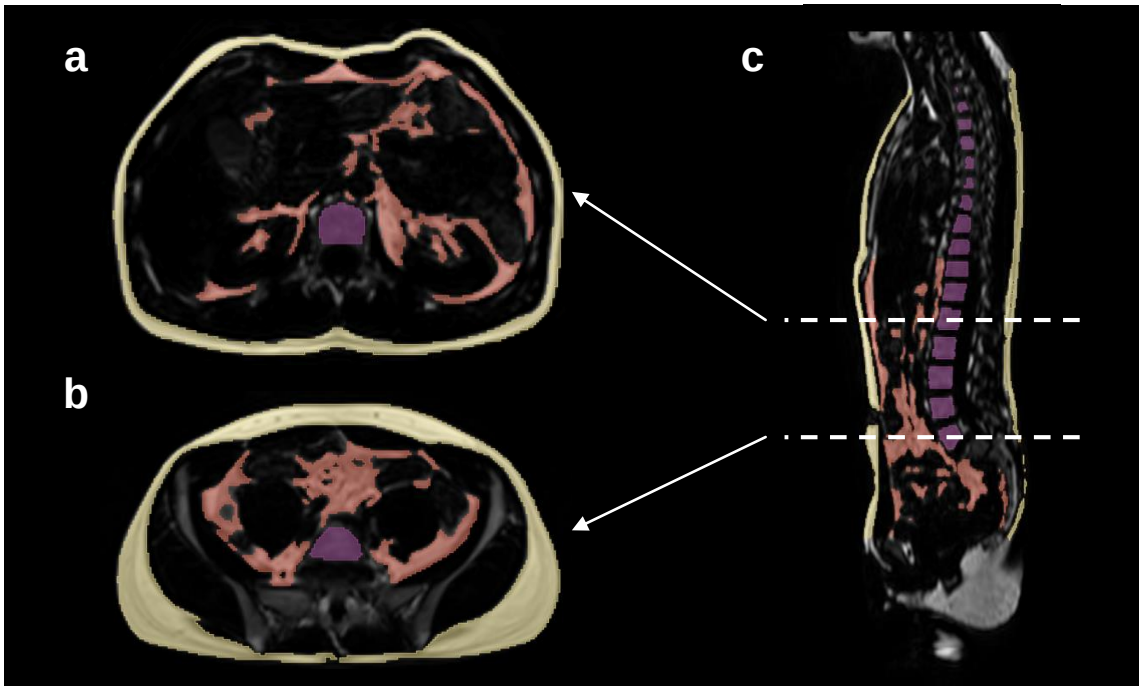

Figure 1.1: **Localization of adipose tissue compartments.** Differentiation of visceral (red) and subcutaneous (yellow) adipose tissue in the trunk at the level of lumbar vertebrae L1 (a) and L5 (b) by semantic image segmentation on a fat-selective Dixon MRI of a 31-year-old man. Intraabdominal and intrapelvic adipose tissue depots are combined to visceral adipose tissue. Localization of the spine (purple) is used to match axial slices to distinct anatomical locations.

Besides cross-sectional and/or discrete cluster analysis, detailed adipose tissue distribution is also used for clustering of longitudinal risk trajectories of cardiometabolic diseases (Rospleszcz et al. 2019; Niedermayer et al. 2024) or analysis of disease heterogeneity, for example of type 2 diabetes, on continuous scale tree-like graph representations (Schön et al. 2024). Additionally, increase of BMAT, for example with aging, is likely to affect systemic disorders, for example, in vertebral bone marrow: bone regeneration can be impaired due to excess fat accumulation (Ambrosi et al. 2017), functionality and support of mesenchymal stem cells in hematopoiesis can be altered (Pachón-Peña et al. 2022). Furthermore, there is an increased risk of osteoporosis and vertebral fractures as increased BMAT is associated with lower bone mineral density (Jung et al. 2023; Beekman et al. 2023). Further, PDFF, the corresponding imaging biomarker, and changes in triglyceride composition assessed by localized single-voxel magnetic resonance spectroscopy (^{1}H-MRS) have been associated with an increased risk of fractures independent of bone mineral density (Beekman et al. 2023).

Application of advanced body composition analysis to large cohort studies

holds potential to advance the understanding of the role of adipose tissue distribution on the development and states of metabolic diseases. Additionally, there is the potential to illustrate the prevalence and associated cardiometabolic disease burden of adipose tissue compartments on a population-based level. Advanced body composition analysis may also help to identify potential regional, age- or sex-specific characteristics, which might indicate the need for localized, personalized specific public health interventions. Moreover, MR-derived phenotypes of adipose tissue distribution can be used to complement existing risk prediction models and will improve individualized risk estimation. Eventually, this will inform clinical decision making, as earlier identification of individuals at risk will lead to more timely and individualized treatment.

After this general introduction to the medical background, the history and applications of body composition analysis, the upcoming sections will delve into greater detail on the assessment of adipose tissue compartments using MRI (sec. 1.1), on automated quantification methods (sec. 1.2), and on the cohort data available (sec. 1.3). Based on the material presented, the research questions and contributions of this thesis are formulated (sec. 1.4).

1.1 Assessment of adipose tissue compartments by magnetic resonance imaging

Magnetic resonance imaging (MRI) makes use of the magnetization properties of atomic nuclei by first aligning the protons to a strong, uniform external magnetic field. By irradiating additional energy using radiofrequency pulses at the corresponding resonance frequency (depending on and perpendicular to the external magnetic field), the protons enter an excited state, and their initial longitudinal magnetization gets tilted to the transverse plane. During relaxation to the equilibrium, an induced voltage signal is measured using receiving coils (Kwok 2022). Tissues can be characterized by two different relaxation times: T1 (longitudinal relaxation time) and T2 (transverse relaxation time). T1 measures the rate at which excited protons return to equilibrium and realign with the external magnetic field, while T2 measures the rate at which excited protons lose phase coherence among the nuclei spinning perpendicular to the main field. Adipose tissue (AT) and lean tissues, for example, muscles or organs, are characterized by distinct T1 and T2 values. Although T1 and T2 generally differ between different organs, they are relatively constant for AT across different AT compartments.

As reviewed by Bojorquez et al., relaxation times vary with respect to 1) the external magnetic field strength, 2) inhomogeneities and/or non-uniformity of the transmitted radiofrequency, 3) measurement methods, as reflected in contradictions of reported values, "possibly because of inconsistent and different measurement methods" (Bojorquez et al. 2017), 4) noise, "generally not considered" in models fitting the received signal (Bojorquez et al. 2017) and 5) tissue pathologies, as reported values generally stem from healthy individuals (Bojorquez et al. 2017). Stikov et al. also reported poor agreement between different methods for the measurement of T1, "with the white matter T1 histogram peak in healthy subjects varying by more than 30% depending on the method used" (Stikov et al. 2015). In general, AT has the lowest T1 making it appear brighter than lean tissues in T1-weighted images. Application of T1-weighted sequences, using either gradient echo imaging or (fast) spin-echo imaging techniques, are preferred for capturing multiple slices quickly and effectively, separating AT from lean tissue due to their shorter T1 relaxation times (Machann et al. 2013).

In the framework of the *Tuebingen Diabetes Family Study* (TDFS) standardized data sets of T1-weighted fast spin echo whole-body MRI have been acquired in more than 1,000 participants at 1.5 Tesla. In the *Tuebingen Lifestyle Intervention Program* (TULIP) study, a longitudinal intervention study examining people at risk to develop type 2 diabetes (T2D), SAT and VAT of 400 participants has been quantified and standardized characterization of AT distribution along the craniocaudal body axis in the form of AT profiles was first taken into account (Machann et al. 2005b). As part of the anatomical standardization, Machann et al. defined VAT as the sum of intrapelvic and intraabdominal AT, measured from the femoral heads to the cardiac apex (Machann et al. 2005b). In follow-up, it was found that VAT and liver fat could be significantly reduced during lifestyle intervention (Machann et al. 2010).

Dixon MRI leverages the chemical shift, which is the difference in resonance frequencies of hydrogen protons depending on their chemical bonds. This results in a frequency difference of about 430 Hz between water protons and methylene protons (the dominant fat peak) at a magnetic field strength of 3 Tesla. This frequency difference leads to distinct phase differences between water and fat protons, which can be used in MRI. In particular, a 180-degree phase shift occurs approximately every 1.23 ms at 3 Tesla, causing signals of water and fat protons to alternate between in-phase (IP) and opposed-phase (OP) relationships. By acquiring IP images (containing the sum of the signal contributions of water and fat protons), for example, at an echo time of 2.46 ms and OP images (containing

the difference of the signal contributions of water and fat protons), for instance, at an echo time of 1.23 ms, water- and fat-selective images can be created by subtraction or addition (Eggers et al. 2014) (see Fig. 1.2). Besides this ability to distinguish between water and fat content within different types of tissue, even in regions with inhomogeneous magnetic field distribution, the technique can also be used to measure fat fraction (FF) at the voxel level (Dixon 1984; J Ma 2008; Eggers et al. 2014).

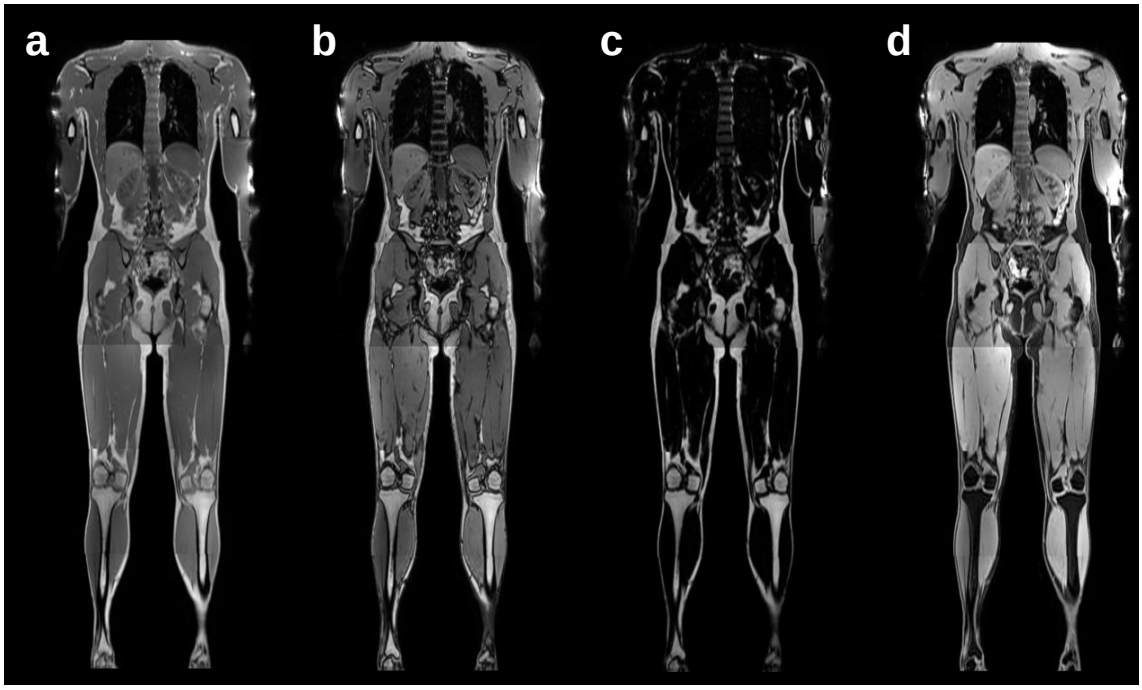

Figure 1.2: **Example of two-point Dixon MRI.** Whole-body two-point Dixon MRI of a 31-year-old man. In-phase (a) and opposed-phase (b) images are acquired and used to calculate fat- and water-selective images. The fat-selective image (c) results from the subtraction of in- and opposed-phase images, in contrast to the water-selective image (d) resulting from addition.

In contrast to original two-point Dixon imaging, which acquires images at two echo times (for example, IP and OP) but requires an internal reference for (magnitude-based) FF estimation, multi-echo (usually using six echoes) Dixon imaging addresses several confounding factors, i.e., T1 bias, (effective) T2* decay, spectral complexity of fat, phase errors, and noise (Karampinos et al. 2018). For example, to handle the spectral complexity of triglycerides, a precalibrated multi-peak fat spectrum, typically derived from single-voxel ^{1}H-MRS measurements of the liver, is included into the signal model (Reeder et al. 2010; Bashir et al. 2015). T1 bias is minimized by selecting a small flip angle for excitation of the magnetization. Presence of iron due to paramagnetic hematopoietic cells in vertebral

bone marrow or iron overload in liver pathologies and presence of trabecular bone leads to inhomogeneities of the magnetic field, resulting in rapid signal decay and shortened T2* corrupting estimates of FF. This confounding effect needs distinct correction during postprocessing or incorporation of T2* into the signal model (Sollmann et al. 2021). After correction for confounding factors, the "density of hydrogen protons attributable to fat, or the fraction of 'unconfounded' proton signal from mobile fat, normalized by the total hydrogen proton density from all mobile proton species" (Reeder et al. 2010) (i.e., proton density fat fraction, PDFF) allows for absolute quantification of FF. Initially developed for PDFF quantification in the liver and "now established as the most accurate and precise method" (Starekova et al. 2021), imaging-based assessment of PDFF is also of interest for other organs and tissues, for example, in skeletal muscles, bone marrow, or in the pancreas.

Applied as volumetric interpolated breath-hold examination (VIBE), Dixon MRI is suitable for whole-body imaging: on a general population level, the *Study of Health In Pomerania* (SHIP) first implemented MRI in a large cohort in Germany: between 2008 and 2012, VIBE T1-weighted imaging was acquired in more than 3,500 study participants at 1.5 Tesla (Hegenscheid et al. 2009; Völzke et al. 2011). The German cohort of the *European Prospective Investigation into Cancer and Nutrition* (EPIC) was the first study to acquire VIBE two-point Dixon imaging at 1.5 Tesla on a population-based level (Wald et al. 2012; Neamat-Allah et al. 2014). Within the *Cooperative Health Research in the Region of Augsburg* (KORA FF4) cohort, VIBE two-point Dixon imaging was first implemented at 3 Tesla in 400 participants as part of a whole-body MRI assessment (Bamberg et al. 2016). Outside of Germany, population level imaging at a very large scale is implemented in the UK Biobank (UKB), aiming at acquiring, besides brain and cardiac MRI, whole-body (head to feet) VIBE two-point Dixon at 1.5 Tesla in 100,000 individuals (Littlejohns et al. 2020). By mid-2024, MRI data from more than 85,000 participants have been acquired.

1.2 Localization and quantification of adipose tissue compartments

Localization and quantification of AT compartments, subsumed under to term 'assessment', i.e., estimating the volume of different depots, is based on the identification of regions of interest, for example, SAT and VAT by semantic image segmentation, i.e., assigning a category to each voxel of an MR image (see Fig. 1.1), followed by voxel counting. Counting uses the property of the MR images that voxels containing AT are brighter compared to voxels not containing AT. The relationship between the number of voxels containing AT and the volume is established by multiplication with the known voxel dimensionality (in-plane pixel dimensions and the slice/partition thickness in 2D/3D MRI). Which AT compartments can be evaluated depends mainly on positioning of research participants. For example, AT in the upper extremities can be assessed in prone position with arms straight above the head (Thomas et al. 1998; Machann et al. 2005b; Thomas et al. 2012), whereas in supine position with arms against the body, the upper extremities usually are excluded. Machann et al. characterized insulin resistant participants by lower percentage of AT in upper extremities compared to insulin sensitive participants (Machann et al. 2005b). However, their results were preliminary due to a very small sample size. There seems to be no obvious connection between AT in the upper extremities as few other studies have considered AT in the upper extremities. For example, Gallagher et al. found no significant differences in arm AT by diabetes status (Gallagher et al. 2009). Recent large cohort studies acquire MRI in supine position, thereby excluding the analysis of upper extremity AT by design.

Available methods for image segmentation developed over time and changed in their approach from manual and semiautomatic to fully automated methods. Crucial to all methods, is the anatomical standardization to obtain more objective AT measurements, particularly regarding separation of VAT. Thomas et al. considered internal fat from the femoral heads to the slice containing the top of the liver or the base of the lungs as visceral (Thomas et al. 1998). Another possibility is taking the cardiac apex as a surrogate measure for the diaphragm which is not visible on MRI [Machann2005]. In the past decades, automatization of the AT quantification implemented histogram-/thresholding-based approaches in combination with manual drawing (Machann et al. 2005b), inclusion of prior anatomical knowledge and region-growing (Thomas et al. 1998), fuzzy clustering in combination with region-growing (Würslin et al. 2010; Müller et al. 2011), statistical shape

models (Wald et al. 2012) or (multi) atlas-based approaches (Borga et al. 2015). More recently, with introduction of the seminal U-Net architecture Ronneberger et al. 2015; Çiçek et al. 2016, deep learning (DL)-based methods have been proposed for semantic AT segmentation (Estrada et al. 2020; Küstner et al. 2020; Bhanu et al. 2022). A more general perspective on DL-based medical image segmentation and a state-of-the-art implementation for many different applications was introduced with nnU-Net that automated not only the image segmentation itself, but also the design of the entire segmentation pipeline including topological parameters of the model architecture as well as image pre- and postprocessing (Isensee et al. 2021). Instead of developing a specialized method for every new problem, the automatic adaptation capabilities of nnU-Net to the characteristics of a given data set substantially simplified the error-prone and time-consuming experimentation process of DL model optimization.

1.3 Cohort data

For the scope of this thesis, MRI data from two different cohorts of healthy subjects are considered: The German National Cohort (*NAKO Gesundheitsstudie*, NAKO) and prospective intervention studies in the framework of the German Center for Diabetes Research (*Deutsches Zentrum für Diabetesforschung*, DZD).

1.3.1 German National Cohort

NAKO is a multidisciplinary, population-based, longitudinal cohort study that aims to investigate the causes of widespread diseases, identify and better characterize lifestyle and socioeconomic risk factors, and improve early detection and prevention of major chronic diseases such as cardiovascular diseases, cancer, diabetes, neurodegenerative and psychiatric diseases, musculoskeletal diseases, and respiratory and infectious diseases. Between 2014 and 2019, over 200,000 participants aged 19 to 74 years were randomly selected and examined across 18 study centers. NAKO collects follow-up information on incident diseases through a combination of active follow-up via written questionnaires at 2–3-year intervals and passive follow-up via record linkages. All participants are invited for re-examinations every four to five years to gather longitudinal data on changes in risk factor profiles and various health functions (Peters et al. 2022).

For a subset of approximately 30,000 participants, whole-body MRI examinations have been conducted at five imaging sites across Germany (Augsburg,

Berlin, Essen, Heidelberg, and Neubrandenburg) using a dedicated neurologic (morphologic and functional imaging of the brain), cardiovascular (morphologic and functional imaging of the cardiovascular system), thoracoabdominal, and musculoskeletal imaging protocol on standardized 3 Tesla whole-body scanners (Magnetom Skyra, Siemens Healthcare, Erlangen, Germany). For thoracoabdominal imaging, a 3D VIBE two-point Dixon sequence is applied from head to knee (Bamberg et al. 2015) (in contrast to UKB which included Dixon MRI from head to feet). Details on the protocol are given in paper 1 (sec. 2.1). In July 2024, the first MRI examinations of the second follow-up have been conducted.

1.3.2 German Center for Diabetes Research

Ongoing cross-sectional and interventional studies involving metabolic imaging conducted within the framework of the DZD for the study of prevention and remission of T2D. Adaptions of lifestyle and lifestyle interventions, for example, modifications of diet and physical exercise, are standard procedures in the prevention of T2D overarchingly aiming at the loss of body weight ($\geq$7% for people with prediabetes according to current standards of the American Diabetes Association). Within TULIP or the multicentric *Prediabetes Lifestyle Intervention Study* (PLIS), a combined approach of reduced energy intake from fat and saturated fatty acids with increased dietary fiber intake accompanied by advised physical exercise of three hours weekly over the course of one year (PLIS) or two years (TULIP) is recommended. PLIS also included an intensified physical exercise treatment of six hours per week, leading to more pronounced beneficial effects on BMI, insulin sensitivity, and liver fat content (Fritsche et al. 2021). Sandforth et al. showed that detailed AT monitoring can be used to study differences in participants with successful remission of prediabetes compared to non-responding participants: responders to remission of prediabetes were characterized by reduced VAT (Sandforth et al. 2023). Other studies, for example, the DiRECT trial based in Newcastle, UK, focused on nutrition and implemented a total diet replacement (Lean et al. 2018), showing that 10% of the original intervention group were still in remission at year five (Lean et al. 2024).

The imaging protocol of current DZD studies has adopted and extended the thoracoabdominal MRI protocol of the NAKO: MR examinations applying 3D VIBE two-point Dixon and multi-echo Dixon techniques were performed on a 3 T whole-body scanner (Magnetom Vida, Siemens Healthcare, Erlangen, Germany). Subjects were positioned head first in supine position on a 24-channel table-integrated

spine-array coil. For homogeneous coverage of the body trunk, two 18-channel body-array coils were placed on chest and lower abdomen. Multi-echo Dixon was acquired in three axial slabs covering the trunk. Details on the protocol are given in paper 2 (sec. 2.2) and paper 3 (sec. 2.3).

1.4 Objectives of the thesis

Retrospective analysis of large cohort studies as well as consecutive assessment of adipose tissue compartments from MRI in ongoing cross-sectional and interventional studies depend on reliable, precise, objective, and time-saving methods for image analysis. The publications comprising this thesis contribute to the standardization of MR-based localization and quantification of adipose tissue compartments in cohorts of healthy subjects by

1. reliable and precise quantification of visceral and subcutaneous adipose tissue in the trunk, standardized between the femoral and humeral heads, based on data from more than 11,000 participants of the NAKO by development and evaluation of an automated deep learning-based segmentation model.

2. enabling standardized assessment of subregional distribution of vertebral proton density fat fraction in lumbar and thoracic vertebral bodies by development of an automated deep learning-based segmentation model

3. estimation of proton density fat fraction from confounded fat fractions in the liver, skeletal muscles, and vertebral bodies with high reliability by derivation of linear correction equations from matching two-point and multi-echo Dixon images of 100 study participants.

2 Results and discussion

2.1 Analysis of volume and topography of adipose tissue in the trunk: Results of MRI of 11,141 participants in the German National Cohort

Haueise T, Schick F, Stefan N, Schlett CL, Weiss JB, Nattenmüller J, Göbel-Guéniot K, Norajitra T, Nonnenmacher T, Kauczor HU, Maier-Hein KH, Niendorf T, Pischon T, Jöckel KH, Umutlu L, Peters A, Rospleszcz S, Kröncke T, Hosten N, Völzke H, Krist L, Willich SN, Bamberg F, Machann J.
Analysis of volume and topography of adipose tissue in the trunk: Results of MRI of 11,141 participants in the German National Cohort.
Sci Adv. 2023;9(19):eadd0433.

Methodological background and approach
As introduced in sec. 1.2, nnU-Net marked a paradigm shift in the development and application of DL-based segmentation models to a variety of tasks facilitating iterative task-specific design and configuration choices, for example, model architecture, training schedule, or image preprocessing. An additional feature of nnU-Net is its data efficiency, which makes it particularly suitable for applications with limited training data (Isensee et al. 2021). For segmentation of VAT and SAT, training data need to be generated in a time-consuming process by manually annotating on average 180 axial slices (six to eight hours) per data set. Creation of training data was based on the model by Küstner et al. (Küstner et al. 2020) but needed several improvements: 1) stratification of data sets by sex, age, and BMI, 2) removal of (partially visible) arms, 3) clear delineation of abdominopelvic cavity, and 4) removal of vertebral and femoral BMAT.

The network topology of the U-Net model for the segmentation of SAT and VAT, automatically configured by nnU-Net (Isensee et al. 2021), is shown in Fig. 2.1a. For the block of each layer, Fig. 2.1b details the implementation. Exceptions to the default kernel size of 3x3x3 for convolutions or 2x2x2 and 1x1x1 for strides, are marked in Fig. 2.1a. For reasons of clarity, the four additional outputs at each step of the decoder used for deep supervision, i.e., additional auxiliary losses added in the decoder allowing gradients to be injected deeper into the network and facilitating the training of all layers (Isensee et al. 2021), are omitted. Instance

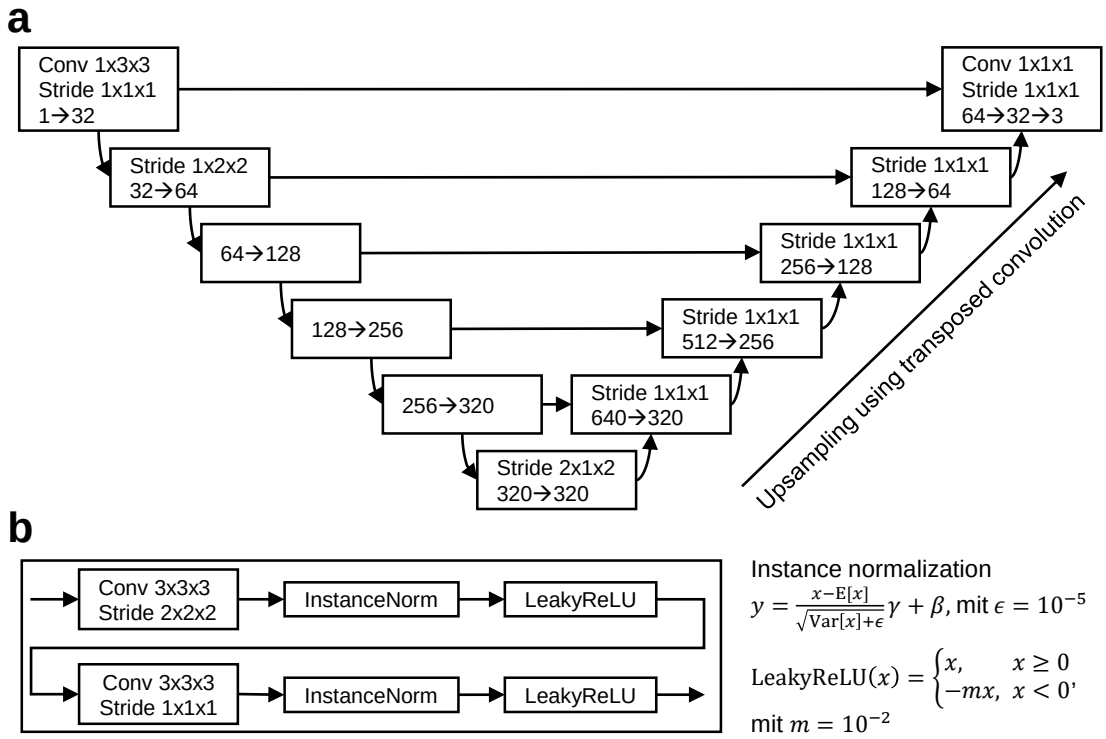

Figure 2.1: **Network topology of nnU-Net model.** Automatically configured U-Net model for the segmentation of VAT and SAT (a) by nnU-Net (Isensee et al. 2021). Each resolution step in (a) consists of the block detailed in (b). Exceptions to default kernel size for convolutions or stride are given in (a). Downsampling from one resolution step to another is implemented as strided convolution. Upsampling in the decoder path is realized by transposed convolutions with strides matching the corresponding downsampling operation in encoder path. Additional outputs for deep supervision are omitted.

normalization is implemented as affine functions and LeakyReLU was used to introduce nonlinearity. The model has a total of 30,758,784 trainable parameters.

nnU-Net prioritizes large patch sizes over large batch sizes, i.e., input patches with a dimensionality of 128x112x128 (Depth×Height×Width) are fed into the network using a batch size of 2. In the encoder path, for each image patch, the number of channels is increased from one (the original grayscale color of the input image) up to 320 in the deepest layer while the image size gets downsampled according to the size of stride kernel, i.e., a stride of 2 bisects the size of the corresponding axis. In the decoder path, the process is reversed using transposed convolutions (Zeiler et al. 2010) for upsampling with strides matching the corresponding downsampling operation in the encoder path. The output convolution with a 1x1x1 kernel follows the default block of the final layer of the decoder. The remaining three channels represent the discrete segmentation mask created by

the model with three classes: background (0), SAT (1), and VAT (2).

Five-fold cross-validated training optimizing the loss function, implemented as the sum of cross-entropy and Dice loss (Drozdzal et al. 2016), of the segmentation model was performed iteratively in two steps, starting with a baseline of 20 randomly selected training data sets with above-mentioned improvements in manual data annotation. After the evaluation of the first nnU-Net training, data sets of five participants with a low BMI of approx. 20 kg/m² per sex were added to the training data set, as the segmentation of fine VAT structures was particularly inaccurate in people with a low BMI.

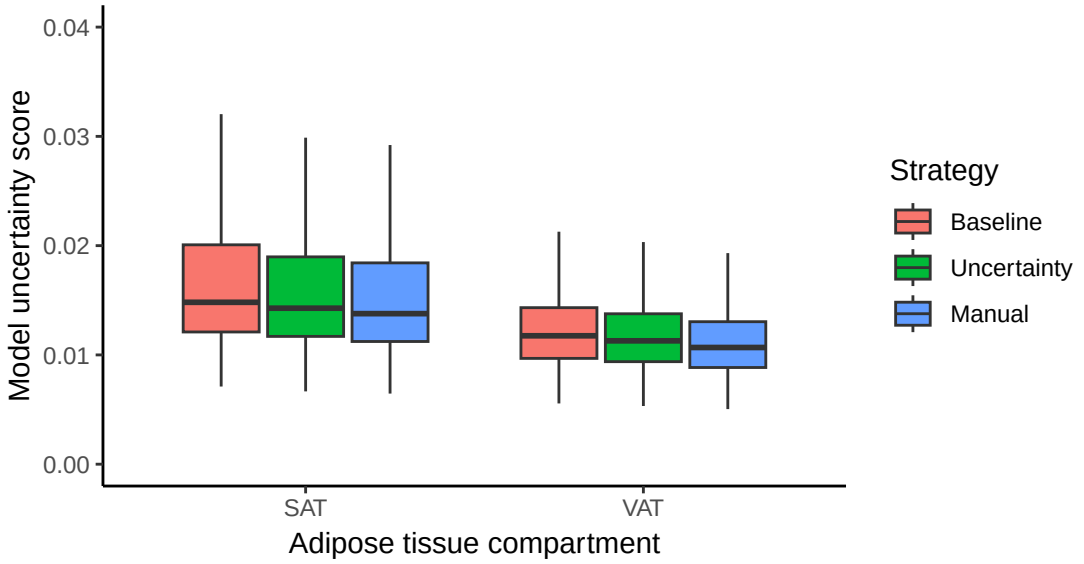

Figure 2.2: **Comparison of two different strategies for training data set extension.** Comparison of the model uncertainty (the lower, the better) for the segmentation of SAT and VAT using two different strategies for the expansion (adding n=10 data sets each, shown in green and blue) of the baseline training data set (n=20, shown in red).

This manual extension strategy based on contextual reasoning was compared with an active learning-inspired metric-based strategy, i.e., adding ten data sets with highest mean model uncertainty. As shown in Fig. 2.2, the model uncertainty score was lower for the training with 10 manually added data sets compared to the metric-based strategy while both models showed improvements compared to baseline. Mean Dice scores were also higher using the manual strategy based on prior knowledge (0.948 for VAT, 0.973 for SAT) compared to the uncertainty-based strategy (0.947 for VAT, 0.971 for SAT) and baseline (0.945 for VAT, 0.970 for SAT).

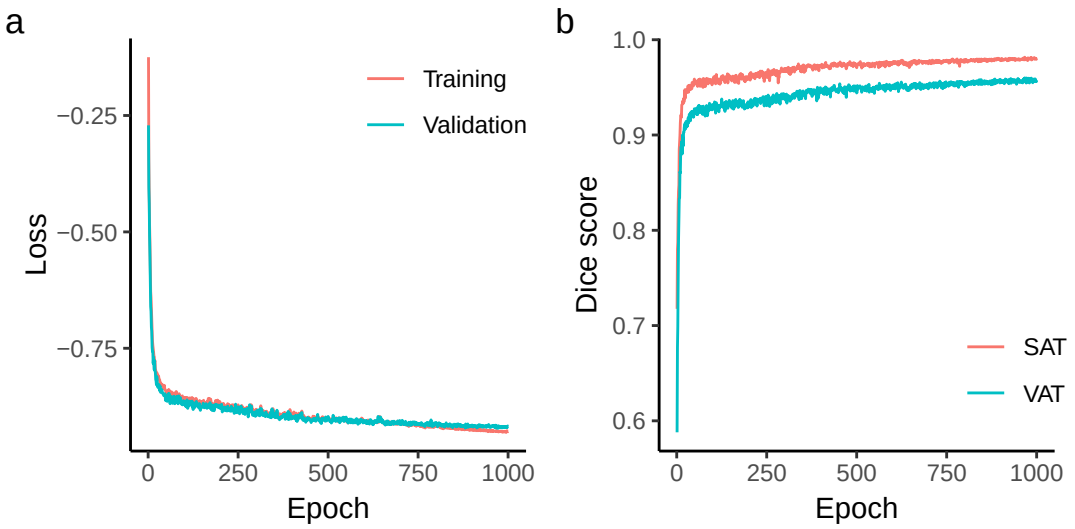

Figure 2.3: **Mean training metrics of nnU-Net model for segmentation of SAT and VAT.** Mean training and validation loss (a) from five-fold cross-validated training and Dice scores for SAT and VAT (b) during training a nnU-Net model for 1000 epochs.

Fig. 2.3 shows the training progress of the nnU-Net model for the segmentation of SAT and VAT in terms of mean loss of training and validation splits (Fig. 2.3a) and Dice score for SAT and VAT as evaluation metric (Fig. 2.3b). Mean duration of one epoch of training was 298 seconds (approx. five minutes), resulting in an average training time per fold of 83 hours on a NVIDIA V100 graphics card.

HEALTH AND MEDICINE

Analysis of volume and topography of adipose tissue in the trunk: Results of MRI of 11,141 participants in the German National Cohort

Tobias Haueise[1,2,3], Fritz Schick[1,2,3], Norbert Stefan[1,2,4], Christopher L. Schlett[5], Jakob B. Weiss[5], Johanna Nattenmüller[5,6], Katharina Göbel-Guéniot[5], Tobias Norajitra[7], Tobias Nonnenmacher[6], Hans-Ulrich Kauczor[6], Klaus H. Maier-Hein[8,9], Thoralf Niendorf[10,11], Tobias Pischon[12,13,14,15], Karl-Heinz Jöckel[16], Lale Umutlu[17], Annette Peters[18,19,20,21], Susanne Rospleszcz[18,19,20], Thomas Kröncke[22,23], Norbert Hosten[24], Henry Völzke[25,26], Lilian Krist[27], Stefan N. Willich[27], Fabian Bamberg[5]†, Juergen Machann[1,2,3]*†

This research addresses the assessment of adipose tissue (AT) and spatial distribution of visceral (VAT) and subcutaneous fat (SAT) in the trunk from standardized magnetic resonance imaging at 3 T, thereby demonstrating the feasibility of deep learning (DL)–based image segmentation in a large population-based cohort in Germany (five sites). Volume and distribution of AT play an essential role in the pathogenesis of insulin resistance, a risk factor of developing metabolic/cardiovascular diseases. Cross-validated training of the DL-segmentation model led to a mean Dice similarity coefficient of >0.94, corresponding to a mean absolute volume deviation of about 22 ml. SAT is significantly increased in women compared to men, whereas VAT is increased in males. Spatial distribution shows age- and body mass index–related displacements. DL-based image segmentation provides robust and fast quantification of AT (≈15 s per dataset versus 3 to 4 hours for manual processing) and assessment of its spatial distribution from magnetic resonance images in large cohort studies.

INTRODUCTION

The obesity pandemic is growing rapidly; in 2016, 39% of the adult world population was overweight, and 13% were obese. The worldwide prevalence has nearly tripled since 1975 (*1*). Abdominal obesity, as manifested by increased visceral adipose tissue (VAT) (*2*), shows a strong correlation to insulin resistance and is a key condition of the metabolic syndrome, which is associated with the risk of developing type 2 diabetes (*3*–*5*) and a major risk factor for a wide range of other diseases (*6*, *7*) such as cardiovascular diseases (*8*, *9*) and several types of cancers (*10*, *11*).

Not only the volume of adipose tissue (AT) but also its regional distribution are considered to play an essential role in the pathogenesis of insulin resistance (*12*, *13*), implying the necessity to characterize individuals for body fat distribution in addition to exclusively determine simple anthropometric measures as, e.g., body mass index (BMI) or waist-to-hip ratio, as, especially VAT shows a better correlation to metabolic parameters (*9*, *14*–*17*). For example, regarding metabolically healthy obesity, in the Tübingen Diabetes Family Study (TDFS), the metabolically healthy and insulin-sensitive obese individuals were found to differ in liver fat content, intramyocellular lipids, and VAT but not in body weight, height, or waist circumference (WC), from the metabolically unhealthy and insulin-resistant obese individuals (*14*). Furthermore, in the TDFS, insulin secretion failure, insulin resistance, fatty liver [measured by ^{1}H magnetic resonance (MR) spectroscopy], and MR imaging (MRI)–determined visceral obesity, but not BMI categories or visceral obesity based on WC measurement, were independent determinants of prediabetes (*18*). Therefore, noninvasive assessment using whole-body MRI, which is able to precisely distinguish between VAT and subcutaneous adipose tissue (SAT), has been established (*19*) and can be regarded as gold standard for the assessment of topography and quantification of AT. State-of-the-art MRI techniques enable gapless acquisitions with high spatial resolution as provided by three-dimensional (3D) chemical shift selective MRI using Dixon-based techniques (*20*, *21*). Large population-based cohort studies such as the German National Cohort (GNC) (*22*) or the U.K. Biobank (*23*) provide comprehensive databases for the assessment of AT depots from MRI (*24*).

Volumetric localization and quantification of AT from MRI are based on slice-wise semantic segmentation of AT compartments. Manual segmentation requires trained personnel, is time-consuming, is costly and—especially in large cohort studies using whole-body images—not feasible in practice. Recent studies have implemented automated segmentation algorithms using atlas-based segmentation (*21*, *25*, *26*), statistical shape models (*27*, *28*), or machine learning (*29*–*31*) on 2D or 3D data using 2D and 3D segmentation algorithms (*32*) and demonstrated the applicability of the methods in small- to medium-sized populations.

Because of the success of deep learning (DL) algorithms in medical image analysis (*33*), many task-specific and highly specialized DL models, often mainly focused on improving model training evaluation metrics, have been proposed (*30*, *31*, *34*, *35*). Because of specific assumptions on input data and nontrivial, often undocumented configuration, the applicability of most of these models in a broader scientific setting is limited. Furthermore, these models do not necessarily output anatomically accurate results despite improved training evaluation metrics, as time resources are invested in an iterative trial-and-error process during method design instead of providing accurate examples of manual segmentation (model-centric DL). Recently, this issue was addressed by the introduction of nnU-Net (*36*). This framework quantitatively confirms that the configuration (including data processing pipeline, training

parameters, etc.) of a DL model has more impact on its performance than architectural variations. Consequently, nnU-Net enables cross-task generalization and can be used as an out-of-the-box tool (*36*) paving the way toward data-centric artificial intelligence that is focused on applications of DL by improving its underlying data (*37*).

The purpose of this study is to assess the volume of different AT compartments of the body trunk, i.e., VAT ranging from hip to cardiac apex, and SAT, which is differentiated in subcutaneous abdominal adipose tissue (SAAT) and subcutaneous thoracic adipose tissue (STAT) using the cardiac apex as the boundary and their spatial distribution along the craniocaudal axis, thereby demonstrating the feasibility of using DL-based image segmentation in a large population-based cohort undergoing MRI.

RESULTS

Automatic data processing using DL segmentation model

On the basis of 30 stratified randomly selected samples from the GNC, fivefold cross-validated training of the nnU-Net segmentation model (exemplary results shown in Fig. 1) led to mean Dice similarity coefficients (DSCs) for VAT, SAAT, and STAT of 0.947 ± 0.033 (0.855 to 0.983), 0.981 ± 0.011 (0.933 to 0.993), and 0.955 ± 0.028 (0.850 to 0.984), corresponding to a mean absolute volume deviation of AT volume of −18.4, 27.5, and 20.3 ml, respectively. Bland-Altman plots (see Fig. 2) showed good agreement and low bias of manual and automated quantification of AT in all three compartments. Comprehensive cross-validation model performance metrics are summarized in Table 1. Intrareader similarity (IRS) of the main annotator is 0.916 (SAAT), 0.876 (STAT), and 0.777 (VAT).

On the basis of a population of 11,191 participants of the GNC, the application of the trained segmentation model led to the uncertainty-based detection of 217 (about 2% of the entire population) potential outliers. After their manual inspection, 21 participants (about 10% of the automatically initially classified outliers) had to be excluded because of imaging errors (partial fat-water swaps) (see Fig. 3, A and B) in the abdomen. False-positive outliers that could be kept after manual inspection mostly include participants with very low AT volume (see Fig. 3, C and D). Manual inspection of 1120 additional participants was unremarkable. In addition, two participants had to be excluded because of corrupted image data, and 27 had to be excluded because of missing height or weight measurements yielding a total of 11,141 participants (5708 males and 5433 females) for AT quantification to form the study population. WC was available from 11,117 participants (5697 males and 5420 females). Anthropometric data of the analyzed study population can be found in Table 2.

Assessment of AT volume and distribution

Regarding the entire study population, females were characterized by significantly higher SAAT and STAT compared to males (7.68 ± 3.88 and 3.78 ± 1.77 liters for females and 6.17 ± 3.05 and 2.81 ± 1.17 liters for males, respectively; see Fig. 4, first and second column, A, B, D, E, G, and H). Males had significantly higher VAT volume (4.84 ± 2.36 liters for males and 2.51 ± 1.55 liters for females; see Fig. 4, last column, C, F, and I).

Females showed a stronger correlation of SAAT, STAT, and VAT with BMI compared to males. SAAT showed the strongest correlation with BMI in both genders. All correlation coefficients are summarized in Table 3. Moreover, normal-weight individuals of both genders show variability in VAT (0.5 to 9.4 liters for males and 0.2 to 5.9 liters for females) and SAAT (0.7 to 8.9 liters for males and 0.8 to 11.3 liters for females). The range of variability of VAT (1.6 to 15.3 liters for males and 1.2 to 10.3 liters for females) and SAAT (3.4 to 26.3 liters for males and 5.9 to 29.4 liters for females) is even greater for obese individuals indicating the presence of the "thin outside fat inside" phenotype with a high share of VAT despite being lean (*38*) and metabolically healthy obese individuals with a low share of VAT, despite being obese (*14*).

Regarding the association with age, SAAT and STAT showed negligible correlation in both genders. VAT showed a moderate but significant positive correlation with age in both genders (see Table 3). Considering age decades, participants in the oldest group of the study population (age > 60 years) had 3.33 ± 1.60 liters of VAT compared to 1.27 ± 0.75 liters (+162%) in the youngest age group (age < 30 years) for women and 5.84 ± 2.34 liters of VAT

[1]Institute for Diabetes Research and Metabolic Diseases, Helmholtz Center Munich at the University of Tuebingen, Tuebingen, Germany. [2]German Center for Diabetes Research (DZD), Tuebingen, Germany. [3]Section on Experimental Radiology, Department of Diagnostic and Interventional Radiology, University Hospital Tuebingen, Tuebingen, Germany. [4]Department of Internal Medicine, Division of Diabetology, Endocrinology and Nephrology, Eberhard-Karls University Tuebingen, Tuebingen, Germany. [5]Department of Diagnostic and Interventional Radiology, Medical Center–University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [6]Department of Diagnostic and Interventional Radiology, University Hospital Heidelberg, Heidelberg, Germany. [7]Division of Medical and Biological Informatics, German Cancer Research Center, Heidelberg, Germany. [8]Division of Medical Image Computing, German Cancer Research Center, Heidelberg, Germany. [9]Pattern Analysis and Learning Group, Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany. [10]Berlin Ultrahigh Field Facility (B.U.F.F.), Max-Delbrueck Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany. [11]Experimental and Clinical Research Center, A Joint Cooperation Between the Charité Medical Faculty and the Max-Delbrueck Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany. [12]Max-Delbrueck-Center for Molecular Medicine in the Helmholtz Association (MDC), Molecular Epidemiology Research Group, Berlin, Germany. [13]Max-Delbrueck-Center for Molecular Medicine in the Helmholtz Association (MDC), Biobank Technology Platform, Berlin, Germany. [14]Berlin Institute of Health at Charité–Universitätsmedizin Berlin, Core Facility Biobank, Berlin, Germany. [15]Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. [16]Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University Hospital Essen, Essen, Germany. [17]Department of Diagnostic and Interventional Radiology and Neuroradiology, University Hospital Essen, Essen, Germany. [18]Department of Epidemiology, Institute for Medical Information Processing, Biometry and Epidemiology, Ludwig-Maximilians-Universität München, Munich, Germany. [19]Institute of Epidemiology, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany. [20]German Center for Cardiovascular Research (DZHK), Partner Site Munich Heart Alliance, Munich, Germany. [21]German Center for Diabetes Research (DZD), Partner Site Neuherberg, Neuherberg, Germany. [22]Department of Diagnostic and Interventional Radiology, University Hospital Augsburg, Faculty of Medicine, University of Augsburg, Augsburg, Germany. [23]Centre for Advanced Analytics and Predictive Sciences (CAAPS), University Augsburg, Augsburg, Germany. [24]Institute of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany. [25]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. [26]German Centre for Cardiovascular Research (DZHK), Partner Site Greifswald, Greifswald, Germany. [27]Institute of Social Medicine, Epidemiology and Health Economics, Charité–Universitätsmedizin Berlin, Berlin, Germany.
*Corresponding author. Email: juergen.machann@med.uni-tuebingen.de
†These authors contributed equally to this work.

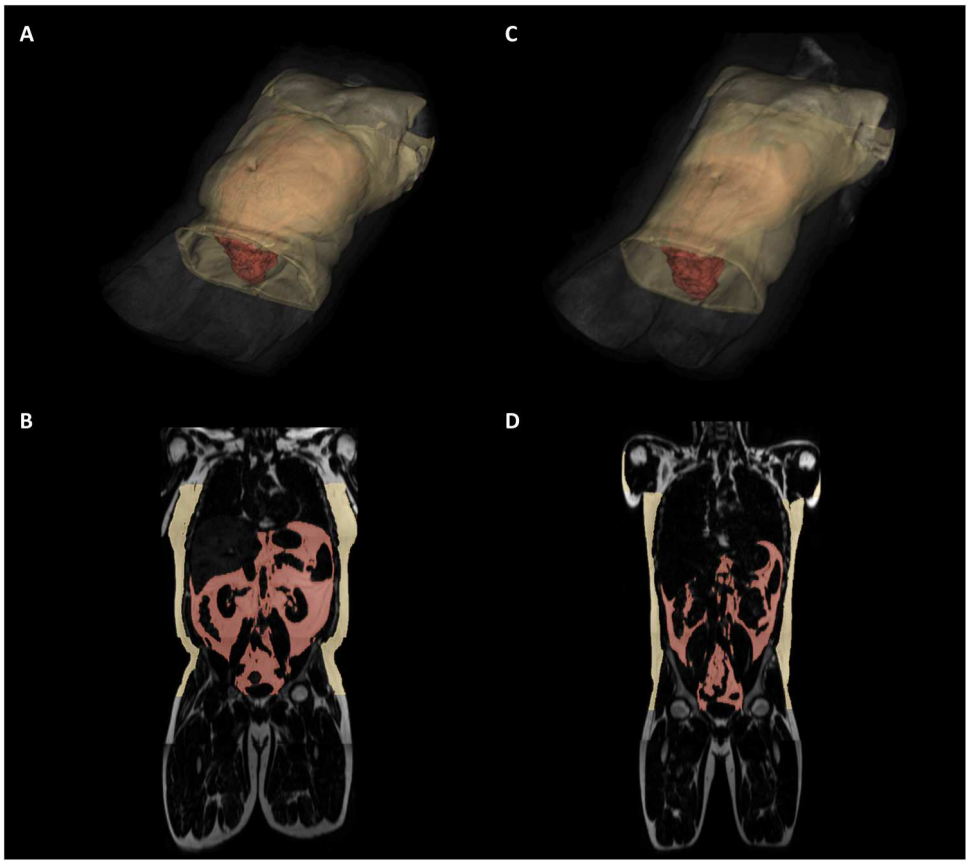

Fig. 1. Examples of successful segmentations. 3D view and exemplary coronal slice of segmented SAT (yellow) and VAT (red). (**A** and **B**) Male participant, 42 years; BMI, 34.5 kg/m^2; SAAT, 10.5 liters; STAT, 4.2 liters; VAT, 6.8 liters. (**C** and **D**) Male participant, 26 years; BMI, 22.9 kg/m^2; SAAT, 5.9 liters; STAT, 2.8 liters; VAT, 2.4 liters.

compared to 2.18 ± 1.30 liters (+167%) for men with each age group showing wide variability (see Fig. 4, G to I).

Using anthropometric measures routinely collected in clinical practice (i.e., age, height, and weight) to explain the variation in MRI-assessed AT compartments showed that the addition of WC lead to an improved prediction for all AT compartments. While VAT showed no gender-specific differences (R^2 = 0.75 for men and women), the variation of both subcutaneous AT compartments was better explained in women (R^2 = 0.90 for SAAT and R^2 = 0.80 for STAT) compared to men (R^2 = 0.81 for SAAT and R^2 = 0.72 for STAT). An overview of all models is provided in table S1.

Regarding the regional spatial distribution of VAT along the craniocaudal axis, there were significant age-dependent differences in the group of normal-weight males. VAT shifts from the pelvis to the abdomen with increasing age. Normal-weight females do not show such displacement (see Fig. 5, A and C). In addition, obese males had less VAT in the pelvis region and showed an age-dependent displacement of VAT towards the lower abdomen. This observation did not apply to females with obesity (see Fig. 5, B and D). Considering SAT, especially in normal-weight females, a similar displacement of AT from the pelvis to the abdomen was observed. Females with obesity did not show any age dependency of the regional distribution of SAT, whereas males with obesity showed a similar redistribution of SAT (see Fig. 6, B to D).

DISCUSSION

Automated MR image segmentation for the analysis of AT compartments of the body trunk using nnU-Net yields state-of-the-art performance without any manual configuration. On the basis of 30 stratified randomly selected and manually annotated samples from the GNC, the model offers robust and fast segmentation performance in terms of low SD in model evaluation metrics (Table 1) and low detection rate of uncertainty-based outliers. Moreover, an improvement of quantitative and qualitative measures compared to the literature is achieved [e.g., increase in mean DSC by 0.02 for SAT and 0.06 for VAT compared to Küstner *et al.* (*30*), respectively]. With regard to the absolute volumetric error in AT quantification, high agreement with the manual segmentation could be achieved narrowing the reported quantification error range (*29*). Only by using DL-based image processing, large data sizes can be handled in a reasonable amount of time. For example, using a trained nnUNet model for the segmentation of data from a single individual takes about 15 s compared to 3 to 4 hours of pure manual segmentation.

Qualitatively, previous studies (*30*, *34*, *35*) revealed weaknesses in the delineation of VAT leading to inaccurate AT quantification, e.g., by including intermuscular fat around the spine, vertebral bone marrow, the skeletal muscles, and parts of the pelvic cavity or by completely ignoring abdominal AT compartments besides SAT. The anatomically standardized segmentation of VAT and SAT

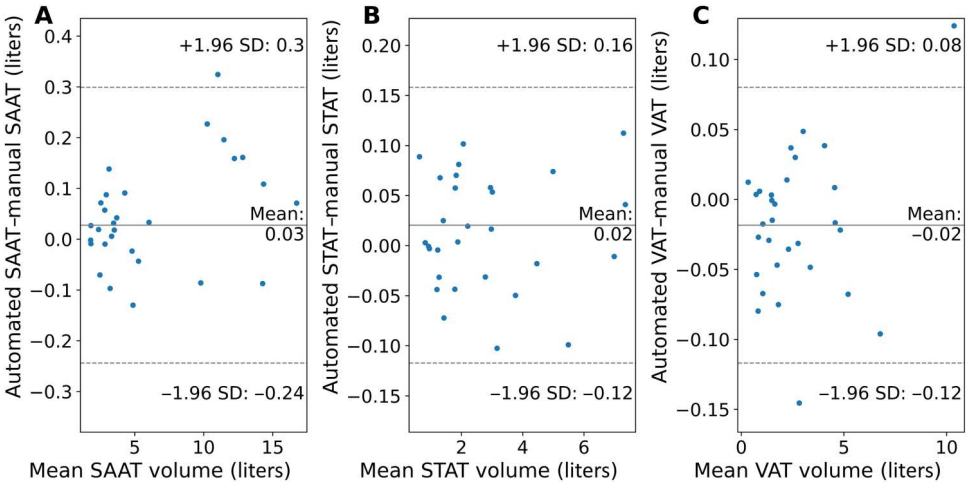

Fig. 2. Agreement of manual and automated segmentation. Bland-Altman plots showing the agreement of SAAT (**A**), STAT (**B**), and VAT (**C**) quantification.

Table 1. Model performance metrics. Mean validation metrics of the class-wise confusion matrix and SD of the fivefold cross-validated model training using a total of 30 annotated datasets.

	SAAT	STAT	VAT
Accuracy	0.998 ± 0.001	0.995 ± 0.002	0.998 ± 0.001
Dice	0.981 ± 0.011	0.955 ± 0.028	0.947 ± 0.033
False discovery rate	0.022 ± 0.014	0.050 ± 0.037	0.047 ± 0.032
False-negative rate	0.017 ± 0.014	0.040 ± 0.027	0.058 ± 0.038
False-omission rate	0.001 ± 0.001	0.002 ± 0.001	0.001 ± 0.001
False-positive rate	0.001 ± 0.001	0.003 ± 0.002	0.001 ± 0.001
Jaccard	0.962 ± 0.027	0.915 ± 0.048	0.902 ± 0.057
Negative predictive value	0.999 ± 0.001	0.998 ± 0.001	0.999 ± 0.001
Precision	0.979 ± 0.015	0.950 ± 0.037	0.953 ± 0.032
Recall	0.983 ± 0.014	0.960 ± 0.027	0.942 ± 0.038
True-negative rate	0.999 ± 0.001	0.997 ± 0.002	0.999 ± 0.001
Relative error (%)	0.52 ± 1.87	1.17 ± 3.94	−1.25 ± 2.79
Absolute error (ml)	27.5 ± 140.9	20.3 ± 71.5	−18.4 ± 51.1

obtained from nnU-Net overcomes these weaknesses. On the basis of unremarkable model uncertainty scores and additional manual review, the model performance directly translates to the large study population of the GNC.

Results from this work are able to confirm and extend the findings of smaller studies in terms of MR population size (*16*, *20*, *39*). First, volumetric segmentation allows assessment of regional spatial AT distribution along the craniocaudal axis leveraging the high spatial resolution of the MR data. A recent study evaluates the association of AT volumes with cardiometabolic diseases but omits the aforementioned advantages using models based on 2D projection images (*40*). Second, increased population size will allow correlations with anthropometric data, age effects, and sex differences in a fine-grained way (*9*).

In the GNC, more than 95% of the participants are Caucasians (*22*), and thus, the presented results most likely reflect this ethnicity with generalizable accuracy. However, because the relations between different AT compartments and their distribution within different ethnic groups are similar within these groups, the described methods are also applicable and can be used to describe other ethnic groups in a similar manner.

Furthermore, MRI-assessed AT compartments, especially VAT, yield deeper insight compared to basic anthropometric measures. Independent of gender, anthropometric measures (age, height, weight, and WC) explain 75% of the variation in VAT corresponding to an estimated standard error of 1.18 liters in men and 0.77 liters in women, respectively. In a study using dual-energy x-ray absorptiometry, it has been reported that approximately 90% of the variation in total fat mass is explained by age, height, weight, and ethnicity (*41*). This study now provides information about the impact of age and simple anthropometric measures on the variation of SAAT, STAT, and VAT (table S1). Together, this study does not intend to abandon any common (and cost effective) measures of abdominal obesity in favor of (expensive) MRI but to show the potential of this technique in terms of differentiation of AT compartments, their volumetric quantification and the possibility of future research of AT distribution along the craniocaudal axis—all of these aspects will probably help in characterizing the general population, taking into account the individual risk for metabolic diseases.

This study has some limitations. First, moderate IRS of the main annotator introduces noise to the training labels of the segmentation model. This noise can lead to systematic errors that are hard to detect as the model is optimized to reproduce the manual segmentations. Second, the automated detection of the region of interest can also introduce small systematic errors by missing VAT depots accumulated close to the diaphragm. Third, this study does not

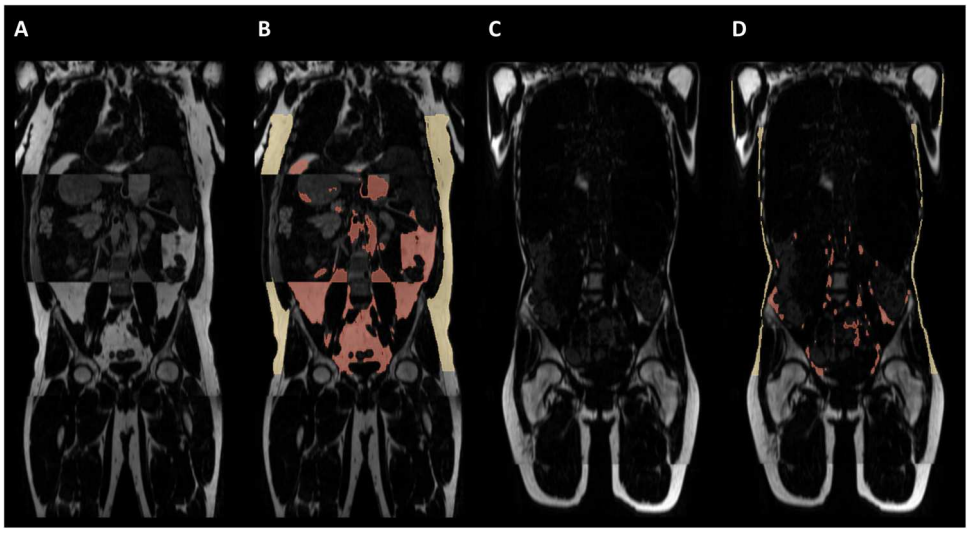

Fig. 3. Outlier detection. Examples of detected outliers based on model uncertainty. (**A** and **B**) True-positive detected outlier due to partial fat-water swap in the abdomen; quantification of AT is not possible [(A) without model–generated segmentation and (B) model-generated segmentation]. (**C** and **D**) False-positive detected outlier due to low AT volume [(C) without model–generated segmentation and (D) model-generated segmentation]; quantification of AT is possible.

Table 2. Study population. Anthropometric data and AT volumes obtained from automatic segmentation of the entire study population and anthropometrics of the training data. *$P < 0.05$.

		Male Means ± SD	Range	Female Means ± SD	Range
Study population	*n*	5708	–	5433	–
	Age (years)	52.1 ± 11.4	20–72	51.7 ± 11.3	20–72
	Height (kg)	178.2 ± 7.0	152.8–204.6	164.9 ± 6.5	126.4–189.0
	Weight (kg)	86.9 ± 14.3	45.8–191.6	71.2 ± 14.4	37.9–157.5
	BMI (kg/m^2)	27.4 ± 4.1	15.6–49.4	26.2 ± 5.2	16.2–54.6
	WC† (cm)	97.4 ± 12.1	63.2–165.0	86.1 ± 13.1	54.0–150.0
	SAAT (liters)	6.17 ± 3.05	0.66–26.3	7.68 ± 3.88	0.83–29.4
	STAT (liters)	2.81 ± 1.17	0.36–13.6	3.78 ± 1.77	0.38–16.0
	VAT (liters)	4.84 ± 2.36	0.54–15.3	2.51 ± 1.55	0.22–10.3
Training data‡	*n*	15	–	15	–
	Age (years)	44.5 ± 14.7*	24–69	44.9 ± 13.4*	23–64
	Height (kg)	175.5 ± 8.0	164.4–190.6	164.5 ± 8.2	152.6–176.5
	Weight (kg)	81.0 ± 15.3	58.7–111.9	66.1 ± 20.0	46.7–106.4
	BMI (kg/m^2)	26.6 ± 6.4	20.0–37.5	24.6 ± 7.9	18.2–37.0
	WC (cm)	91.9 ± 14.6	72.7–125.0	82.0 ± 17.5	64.1–111.6

†Analysis included all individuals with WC data available ($n = 5697$ males and $n = 5420$ females). ‡Differences between the complete study population and the training subset are tested for significance.

evaluate additional variables associated with diseases (metabolic data or laboratory parameters) and is limited to basic anthropometric data and image-based parameters. Fourth, an independent annotated testing dataset is missing and manual inspection of all segmentation results is not feasible because of the size of the population. However, the low true-positive rate (7%) of the outlier detection in combination with a manual check for anomalies (10% of the study population) sufficiently minimizes the probability of unidentified error. Fifth, the applied Dixon technique just allows a binary decision in the presence of fat mass and is not capable to detect small amounts of ectopic fat (e.g., in the liver, pancreas, or skeletal muscles), which would add important additional information on the metabolic condition of the individuals (*14*, *16*, *39*).

In conclusion, the results presented demonstrate the effectiveness of the nnU-net model to provide automated assessment of the volume and topography of AT in humans, with accuracy and

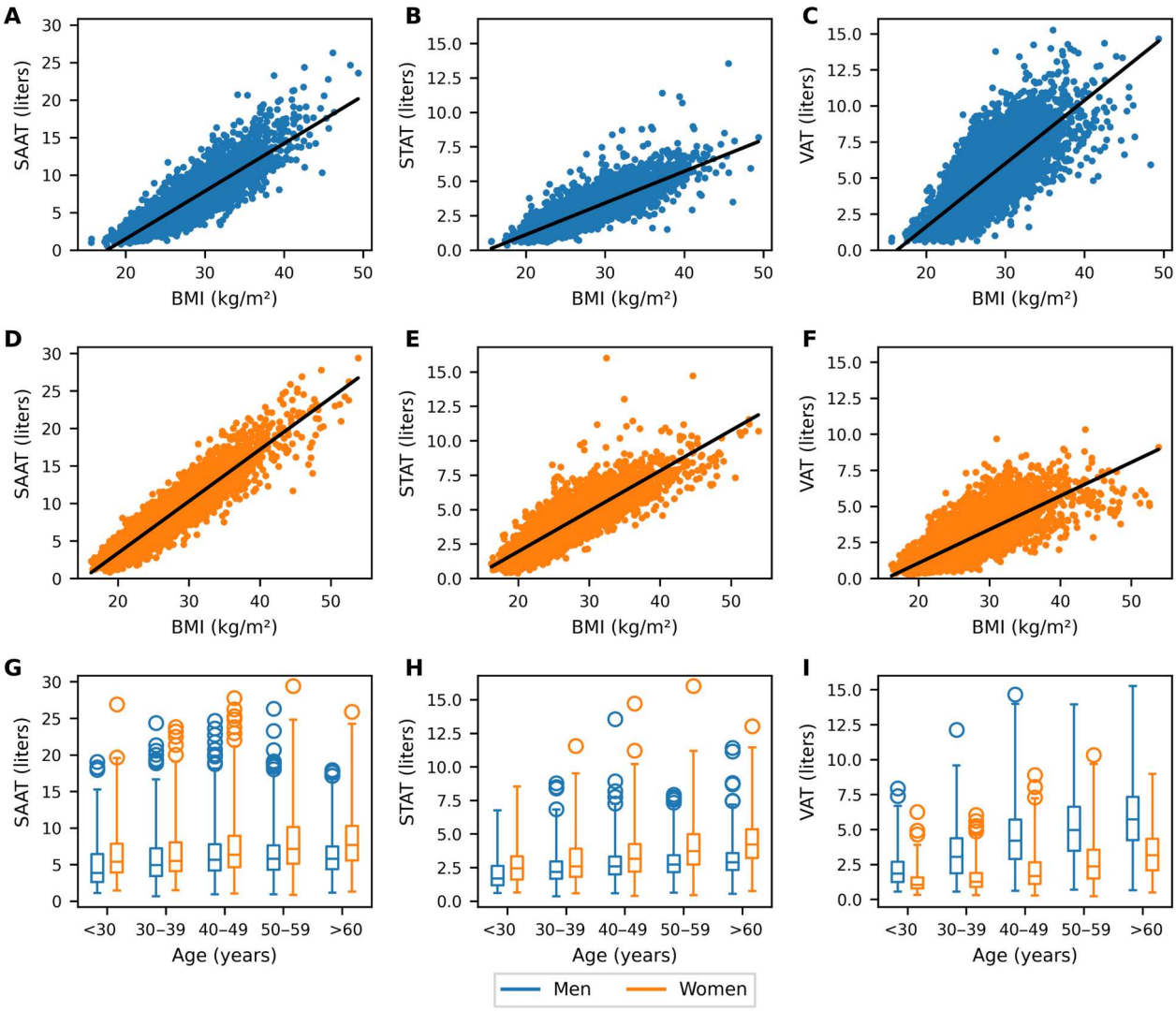

Fig. 4. Anthropometric associations. Association and linear regression of AT compartments with BMI [SAAT for males (**A**) and females (**D**), STAT for males (**B**) and females (**E**), and VAT for males (**C**) and females (**F**)] and with age [SAAT (**G**), STAT (**H**), and VAT (**I**)].

Table 3. Correlations with anthropometric data. *$P < 0.05$ and **$P < 0.001$.

	SAAT		STAT		VAT	
	Male	Female	Male	Female	Male	Female
Age	0.07**	0.15**	0.22**	0.28**	0.40**	0.41**
Height	0.10**	−0.03*	0.05**	−0.10**	−0.01	−0.11**
Weight	0.85**	0.91**	0.78**	0.82**	0.71**	0.74**
BMI	0.86**	0.93**	0.81**	0.87**	0.77**	0.79**
WC†	0.85**	0.89**	0.83**	0.86**	0.85**	0.84**

†Analysis included all individuals with WC data available (n = 5697 males and n = 5420 females).

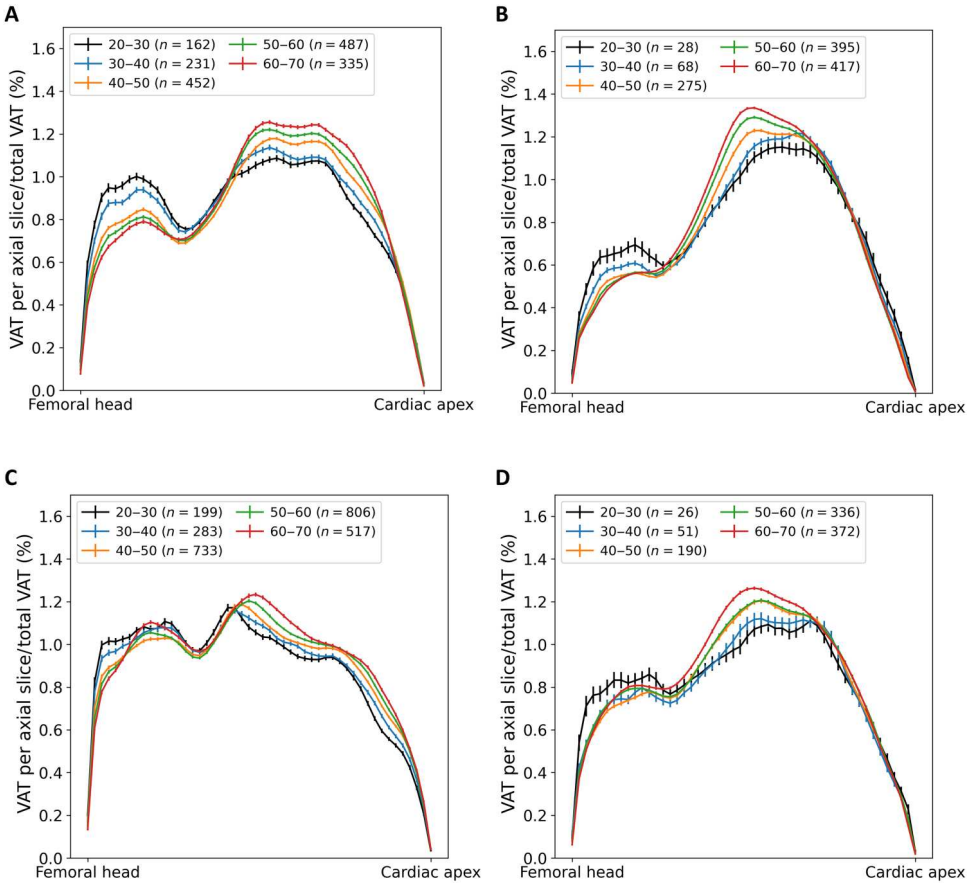

Fig. 5. Spatial distribution of VAT. Age dependency of the regional distribution of VAT along craniocaudal axis. (**A**) Males with a BMI of <25.0 kg/m^2, (**B**) males with a BMI of >30.0 kg/m^2, (**C**) females with a BMI of <25.0 kg/m^2, and (**D**) females with a BMI of >30.0 kg/m^2. Error bars show SEM.

precision equivalent to that of skilled human observers. This approach and the results obtained from the large population of the GNC are relevant for both epidemiological and clinical perspectives. On the basis of automated MR image analysis, meaningful epidemiological data illustrating prevalence and associated cardiometabolic disease burden of AT compartments allow to identify gender-specific and regional characteristics. By complementing existing risk prediction models with characterization of body fat distribution, improved and individualized risk estimation will be possible, as earlier identification of individuals at risk will lead to more timely and individualized prevention and treatment.

MATERIALS AND METHODS

German National Cohort

The GNC (NAKO Gesundheitsstudie) is a population-based, longitudinal multicentric cohort study in Germany enrolling >200,000 participants selected randomly from the population. Its main objective is to identify and to characterize risk factors for major chronic diseases (e.g., diabetes mellitus and cancer) (*22*). For a subset of approximately 30,000 participants, whole-body MRI examinations have been conducted at five imaging sites using dedicated neurologic, cardiovascular, thoracoabdominal, and musculoskeletal imaging protocols (*42*). All local on-site institutional review boards in charge of the five imaging sites approved the GNC, and written informed consent of all participants was obtained before study enrollment.

MRI data acquisition

MRI was performed at five sites using 3-T whole-body scanners (all MAGNETOM Skyra, Siemens Healthineers, Erlangen, Germany) using a standardized acquisition protocol (*42*). Imaging of the body trunk was performed using a dedicated T1-weighted 3D VIBE two-point DIXON sequence in axial orientation with 3 mm in section thickness, 1.4 mm–by–1.4 mm in-plane voxel size, echo times of 1.23 and 2.46 ms, and a repetition time of 4.36 ms according to the GNC protocol (*42*). From this, fat- and water-selective images are automatically calculated on the scanners. For the sake of data minimization, only fat-selective images are used in the analyses. The data used in this study were obtained from the first GNC release of MRI data, which includes 11,191 participants being screened between May 2014 and December 2016.

Segmentation model

For the automated, retrospective analysis of the MR data, a stratified (age and BMI) random sample of 30 (15 males and 15 females; demographics are provided in Table 2) manually segmented fat-selective MR images was used to train a 3D U-Net model (nnU-Net, full-resolution configuration) (*36*) to perform the segmentation of VAT and SAT. The 3D nnU-Net model was trained out of the box for

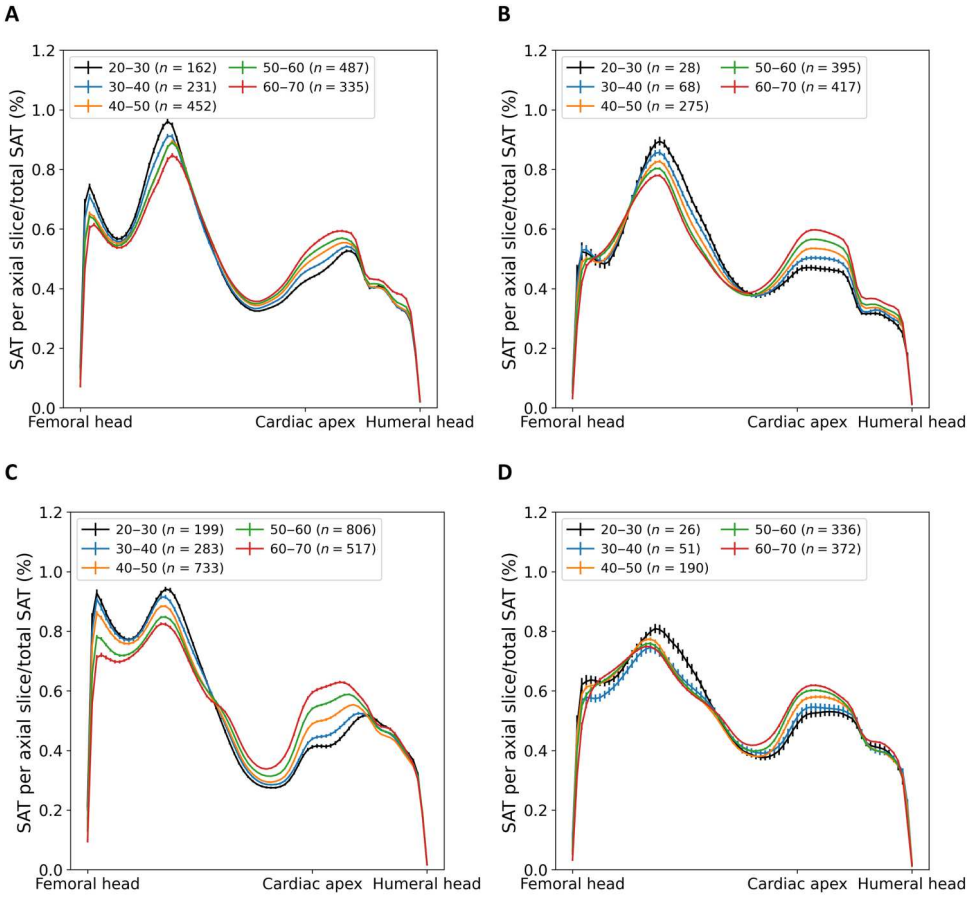

Fig. 6. Spatial distribution of SAT. Age dependency of the regional distribution of SAT along craniocaudal axis. (**A**) Males with a BMI of <25.0 kg/m^2, (**B**) males with a BMI of >30.0 kg/m^2, (**C**) females with a BMI of <25.0 kg/m^2, and (**D**) females with a BMI of > 30.0 kg/m^2]. Error bars show SEM.

1000 epochs using fivefold cross-validation providing the mean (i.e., the output of an ensemble) of the five resulting independent models as resulting segmentation as suggested by the authors (*36*). The model-generated segmentations were evaluated by DSC and by the actual volume of the AT compartments and the percentage of the error.

Manual segmentation

The manual labeling process was performed by a doctoral student under the supervision of two experienced medical physicists. IRS was assessed after a 2-month interruption by resegmentation from scratch of randomly chosen axial slices from the 30 originally annotated datasets.

To assess VAT, defined as AT inside the abdominal cavity including retroperitoneal structures such as the kidneys, pancreas, or duodenum, in a standardized manner, AT accumulated around the heart is excluded. The manual segmentation was performed from the middle of the femoral heads to the cardiac apex, since the thoracic diaphragm cannot be detected on the MR images. SAT was segmented ranging from the middle of the femoral heads to the middle of the humeral heads (see red dashed lines in Fig. 7 indicating the different levels). By design, nnU-Net implicitly recognizes these inferior and superior boundaries with no need for an explicit adaption of the model to the region of interest.

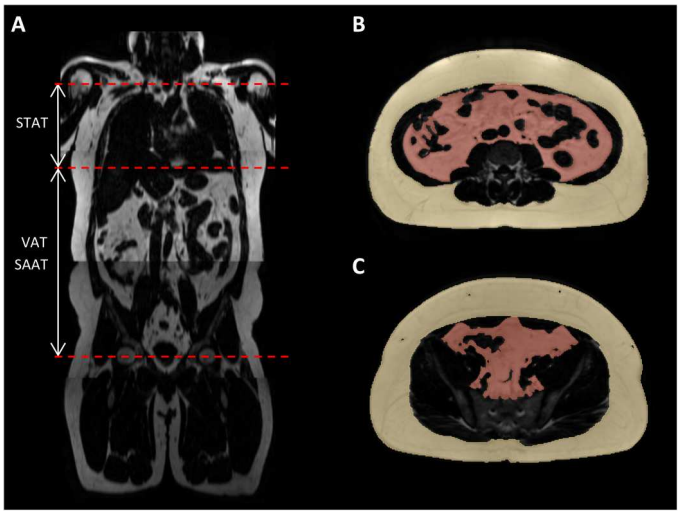

Fig. 7. Manual segmentation. (**A**) Red dashed lines indicating the limits of the ranges of manual segmentation (femoral heads to cardiac apex for VAT and SAAT and cardiac apex to humeral heads for STAT, respectively) and (**B** and **C**) axial examples of manual segmentations of SAT (yellow) and VAT (red).

Uncertainty-based outlier detection

Because of the large population size of the cohort, an automated measure for outlier detection in the segmentation based on mean pairwise DSC of the cross-validated training folds to estimate the model's uncertainty (*43*) was used. Participants are classified as outliers if the model uncertainty is three interquartile ranges below the first quartile of all participants. Detected outliers were reviewed manually by three different scientists. In addition, random datasets up to 10% of the study cohort were manually checked for anomalies.

Fat quantification and spatial distribution

For the analysis of the associations of AT depots and anthropometric data, SAT was differentiated in SAAT and STAT using the cardiac apex as the boundary. The regional spatial distribution of the AT compartments was described by considering the percentage of total AT of the trunk along the craniocaudal axis.

Anthropometric data

Body height and weight were assessed using standardized measuring instruments across the study centers (all Stadiometer 274 for height and medical Body Composition Analyzer 515 for weight, both seca GmBH, Hamburg, Germany). WC was measured at the midpoint between the iliac crest and the lowest rib. The study participants should come to the measurements with an empty bladder, should not have been physically active in the last hour, and should not lie down 10 min before the measurements (*44*). For further analyses, normal weight was defined as BMI $\leq$ 25 kg/m^2, whereas obesity was defined as BMI $\geq$ 30 kg/m^2.

Statistical analysis

Data are reported as means ± SD unless stated otherwise. Bland-Altman plots are were used to visualize the agreement between manual and automated AT quantification. IRS was measured as DSC using true-positive, false-positive, and false-negative annotated image pixels. Model performance was evaluated using class-wise metrics of the confusion matrix, e.g., DSC or precision. The association of AT compartments and anthropometric (gender, age, and BMI) data was assessed using linear regression and Pearson's correlation coefficient. Two-sample Welch's *t* test was used for the determination of gender-related differences. The explained variation in MRI-assessed AT compartments by anthropometric measures is modeled using multiple linear regression models. SEM was used in the visualization of regional AT distribution. $P < 0.05$ was considered statistically significant in this study. All statistical analyses were performed in Python 3.8 using SciPy 1.5.4 and R version 4.2.0.

Supplementary Materials

This PDF file includes:

Table S1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. World Health Organization, Obesity and overweight, in *WHO Fact Sheet* (2021); www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.
2. A. Tchernof, J.-P. Després, Pathophysiology of human visceral obesity: An update. *Physiol. Rev.* **93**, 359–404 (2013).
3. J.-P. Després, I. Lemieux, Abdominal obesity and metabolic syndrome. *Nature* **444**, 881–887 (2006).
4. L. O. Ohlson, B. Larsson, K. Svärdsudd, L. Welin, H. Eriksson, L. Wilhelmsen, P. Björntorp, G. Tibblin, The influence of body fat distribution on the incidence of diabetes mellitus. 13.5 years of follow-up of the participants in the study of men born in 1913. *Diabetes* **34**, 1055–1058 (1985).
5. J.-P. Després, Body fat distribution and risk of cardiovascular disease. *Circulation* **126**, 1301–1313 (2012).
6. G. M. Reaven, Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes* **37**, 1595–1607 (1988).
7. B. L. Wajchenberg, M. Nery, M. R. Cunha, M. E. R. da Silva, Adipose tissue at the crossroads in the development of the metabolic syndrome, inflammation and atherosclerosis. *Arq. Bras. Endocrinol. Metabol.* **53**, 145–150 (2009).
8. C. L. Schlett, R. Lorbeer, C. Arndt, S. Auweter, J. Machann, H. Hetterich, B. Linkohr, W. Rathmann, A. Peters, F. Bamberg, Association between abdominal adiposity and subclinical measures of left-ventricular remodeling in diabetics, prediabetics and normal controls without history of cardiovascular disease as measured by magnetic resonance imaging: Results from the KORA-FF4 Study. *Cardiovasc. Diabetol.* **17**, 88 (2018).
9. I. J. Neeland, R. Ross, J.-P. Després, Y. Matsuzawa, S. Yamashita, I. Shai, J. Seidell, P. Magni, R. D. Santos, B. Arsenault, A. Cuevas, F. B. Hu, B. Griffin, A. Zambon, P. Barter, J.-C. Fruchart, R. H. Eckel, Visceral and ectopic fat, atherosclerosis, and cardiometabolic disease: A position statement. *Lancet Diabetes Endocrinol.* **7**, 715–725 (2019).
10. G. Behrens, T. Gredner, C. Stock, M. F. Leitzmann, H. Brenner, U. Mons, Cancers due to excess weight, low physical activity, and unhealthy diet. *Dtsch. Arzteblatt Int.* **115**, 578–585 (2018).
11. C. L. Donohoe, S. L. Doyle, J. V. Reynolds, Visceral adiposity, insulin resistance and cancer risk. *Diabetol. Metab. Syndr.* **3**, 12 (2011).
12. M. Krotkiewski, P. Björntorp, L. Sjöström, U. Smith, Impact of obesity on metabolism in men and women. Importance of regional adipose tissue distribution. *J. Clin. Invest.* **72**, 1150–1162 (1983).
13. A. H. Kissebah, N. Vydelingum, R. Murray, D. J. Evans, A. J. Hartz, R. K. Kalkhoff, P. W. Adams, Relation of body fat distribution to metabolic complications of obesity. *J. Clin. Endocrinol. Metab.* **54**, 254–260 (1982).
14. N. Stefan, K. Kantartzis, J. Machann, F. Schick, C. Thamer, K. Rittig, B. Balletshofer, F. Machicao, A. Fritsche, H.-U. Häring, Identification and characterization of metabolically benign obesity in humans. *Arch. Intern. Med.* **168**, 1609–1616 (2008).
15. J. Machann, N. Stefan, R. Wagner, A. Fritsche, J. D. Bell, B. Whitcher, H.-U. Häring, A. L. Birkenfeld, K. Nikolaou, F. Schick, E. L. Thomas, Normalized indices derived from visceral adipose mass assessed by magnetic resonance imaging and their correlation with markers for insulin resistance and prediabetes. *Nutrients* **12**, 2064 (2020).
16. J. Linge, M. Borga, J. West, T. Tuthill, M. R. Miller, A. Dumitriu, E. L. Thomas, T. Romu, P. Tunón, J. D. Bell, O. Dahlqvist Leinhard, Body composition profiling in the UK Biobank imaging study. *Obes. Silver Spring Md.* **26**, 1785–1795 (2018).
17. N. Stefan, Causes, consequences, and treatment of metabolically unhealthy fat distribution. *Lancet Diabetes Endocrinol.* **8**, 616–627 (2020).
18. N. Stefan, A. Fritsche, F. Schick, H.-U. Häring, Phenotypes of prediabetes and stratification of cardiometabolic risk. *Lancet Diabetes Endocrinol.* **4**, 789–798 (2016).
19. J. Machann, C. Thamer, B. Schnoedt, M. Haap, H.-U. Haring, C. D. Claussen, M. Stumvoll, A. Fritsche, F. Schick, Standardized assessment of whole body adipose tissue topography by MRI. *J. Magn. Reson. Imaging JMRI.* **21**, 455–462 (2005).
20. F. Bamberg, H. Hetterich, S. Rospleszcz, R. Lorbeer, S. D. Auweter, C. L. Schlett, A. Schafnitzel, C. Bayerl, A. Schindler, T. Saam, K. Müller-Peltzer, W. Sommer, T. Zitzelsberger, J. Machann, M. Ingrisch, S. Selder, W. Rathmann, M. Heier, B. Linkohr, C. Meisinger, C. Weber, B. Ertl-Wagner, S. Massberg, M. F. Reiser, A. Peters, Subclinical disease burden as assessed by whole-body MRI in subjects with prediabetes, subjects with diabetes, and normal control subjects from the general population: The KORA-MRI study. *Diabetes* **66**, 158–169 (2017).
21. J. West, O. Dahlqvist Leinhard, T. Romu, R. Collins, S. Garratt, J. D. Bell, M. Borga, L. Thomas, Feasibility of MR-based body composition analysis in large scale population studies. *PLOS ONE* **11**, e0163332 (2016).
22. W. Ahrens, H. Greiser, J. Linseisen, A. Kluttig, S. Schipf, B. Schmidt, K. Günther, Das design der Machbarkeitsstudien für eine bundesweite Kohortenstudie in Deutschland. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **57**, 1246–1254 (2014).
23. C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, B. Liu, P. Matthews, G. Ong, J. Pell, A. Silman, A. Young, T. Sprosen, T. Peakman, R. Collins, UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Med.* **12**, e1001779 (2015).
24. C. L. Schlett, T. Hendel, S. Weckbach, M. Reiser, H. U. Kauczor, K. Nikolaou, M. Günther, M. Forsting, N. Hosten, H. Völzke, F. Bamberg, Population-based imaging and radiomics:

Rationale and perspective of the German National Cohort MRI study. *ROFO. Fortschr. Geb. Rontgenstr. Nuklearmed.* **188**, 652–661 (2016).
25. A. Karlsson, J. Rosander, T. Romu, J. Tallberg, A. Grönqvist, M. Borga, O. D. Leinhard, Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. *J. Magn. Reson. Imaging JMRI.* **41**, 1558–1569 (2015).
26. J. E. Iglesias, M. R. Sabuncu, Multi-atlas segmentation of biomedical images: A survey. *Med. Image Anal.* **24**, 205–219 (2015).
27. D. Wald, B. Teucher, J. Dinkel, R. Kaaks, S. Delorme, H. Boeing, K. Seidensaal, H.-P. Meinzer, T. Heimann, Automatic quantification of subcutaneous and visceral adipose tissue from whole-body magnetic resonance images suitable for large cohort studies. *J. Magn. Reson. Imaging JMRI.* **36**, 1421–1434 (2012).
28. T. Heimann, H.-P. Meinzer, Statistical shape models for 3D medical image segmentation: A review. *Med. Image Anal.* **13**, 543–563 (2009).
29. T. Langner, A. Hedström, K. Mörwald, D. Weghuber, A. Forslund, P. Bergsten, H. Ahlström, J. Kullberg, Fully convolutional networks for automated segmentation of abdominal adipose tissue depots in multicenter water-fat MRI. *Magn. Reson. Med.* **81**, 2736–2745 (2019).
30. T. Küstner, T. Hepp, M. Fischer, M. Schwartz, A. Fritsche, H.-U. Häring, K. Nikolaou, F. Bamberg, B. Yang, F. Schick, S. Gatidis, J. Machann, Fully automated and standardized segmentation of adipose tissue compartments via deep learning in 3D whole-body MRI of epidemiologic cohort studies. *Radiol. Artif. Intell.* **2**, e200010 (2020).
31. S. Estrada, R. Lu, S. Conjeti, X. Orozco-Ruiz, J. Panos-Willuhn, M. M. B. Breteler, M. Reuter, FatSegNet: A fully automated deep learning pipeline for adipose tissue segmentation on abdominal dixon MRI. *Magn. Reson. Med.* **83**, 1471–1483 (2020).
32. M. Borga, MRI adipose tissue and muscle composition analysis-a review of automation techniques. *Br. J. Radiol.* **91**, 20180252 (2018).
33. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017).
34. P. K. Bhanu, C. S. Arvind, L. Y. Yeow, W. X. Chen, W. S. Lim, C. H. Tan, CAFT: A deep learning-based comprehensive abdominal fat analysis tool for large cohort studies. *Magma N. Y. N.* **35**, 205–220 (2022).
35. Z. Wang, C. Cheng, H. Peng, Y. Qi, Q. Wan, H. Zhou, S. Qu, D. Liang, X. Liu, H. Zheng, C. Zou, Automatic segmentation of whole-body adipose tissue from magnetic resonance fat fraction images based on machine learning. *Magma N. Y. N.* **35**, 193–203 (2022).
36. F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, K. H. Maier-Hein, nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
37. N. Polyzotis, M. Zaharia, What can data-Centric AI learn from data and ML engineering? arXiv:2112.06439v1 [cs.LG] (13 December 2021).
38. E. L. Thomas, J. R. Parkinson, G. S. Frost, A. P. Goldstone, C. J. Doré, J. P. McCarthy, A. L. Collins, J. A. Fitzpatrick, G. Durighel, S. D. Taylor-Robinson, J. D. Bell, The missing risk: MRI and MRS phenotyping of abdominal adiposity and ectopic fat. *Obesity* **20**, 76–87 (2012).
39. J. Machann, C. Thamer, B. Schnoedt, N. Stefan, M. Stumvoll, H.-U. Haring, C. D. Claussen, A. Fritsche, F. Schick, Age and gender related effects on adipose tissue compartments of subjects with increased risk for type 2 diabetes: A whole body MRI/MRS study. *Magma N. Y. N.* **18**, 128–137 (2005).
40. S. Agrawal, M. D. R. Klarqvist, N. Diamant, P. T. Ellinor, N. N. Mehta, A. Philippakis, K. Ng, P. Batra, A. V. Khera, Association of machine learning-derived measures of body fat distribution in >40,000 individuals with cardiometabolic diseases. medRxiv 2021.05.07.21256854 [Preprint]. 28 July 2021. https://doi.org/10.1101/2021.05.07.21256854.
41. D. H. Lee, N. Keum, F. B. Hu, E. J. Orav, E. B. Rimm, Q. Sun, W. C. Willett, E. L. Giovannucci, Development and validation of anthropometric prediction equations for lean body mass, fat mass and percent fat in adults using the National Health and Nutrition Examination Survey (NHANES) 1999–006. *Br. J. Nutr.* **118**, 858–866 (2017).
42. F. Bamberg, H.-U. Kauczor, S. Weckbach, C. L. Schlett, M. Forsting, S. C. Ladd, K. H. Greiser, M.-A. Weber, J. Schulz-Menger, T. Niendorf, T. Pischon, S. Caspers, K. Amunts, K. Berger, R. Bülow, N. Hosten, K. Hegenscheid, T. Kröncke, J. Linseisen, M. Günther, J. G. Hirsch, A. Köhn, T. Hendel, H.-E. Wichmann, B. Schmidt, K.-H. Jöckel, W. Hoffmann, R. Kaaks, M. F. Reiser, H. Völzke; German National Cohort MRI Study Investigators, Whole-body MR imaging in the German National Cohort: Rationale, design, and technical background. *Radiology* **277**, 206–220 (2015).
43. A. G. Roy, S. Conjeti, N. Navab, C. Wachinger, Inherent brain segmentation quality control from fully ConvNet Monte Carlo sampling, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, G. Fichtinger, Eds. (Springer International Publishing, 2018), vol. 11070 of *Lecture Notes in Computer Science*, pp. 664–672; http://link.springer.com/10.1007/978-3-030-00928-1_75.
44. B. Fischer, A. M. Sedlmeier, S. Hartwig, C. L. Schlett, W. Ahrens, F. Bamberg, H. Baurecht, H. Becher, K. Berger, H. Binder, B. Bohn, P. R. Carr, S. Castell, C.-W. Franzke, J. Fricke, S. Gastell, K. H. Greiser, K. Günther, L. Jaeschke, R. Kaaks, Y. Kemmling, L. Krist, O. Kuß, N. Legath, W. Lieb, J. Linseisen, M. Löffler, K. B. Michels, R. Mikolajczyk, T. Niedermaier, K. Norman, N. Obi, A. Peters, T. Pischon, T. Schikowski, S. Schipf, B. Schmidt, M. B. Schulze, A. Stang, J. Stojicic, D. Tiller, H. Völzke, S. Waniek, M. F. Leitzmann, Anthropometric measures in the German National Cohort-more than weight and height. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **63**, 290–300 (2020).

Acknowledgments: We thank all participants who took part in the GNC study and the staff in this research program. We acknowledge support by the Open Access Publishing Fund of the University of Tübingen. **Funding:** This study was funded by German Research Foundation grant 428224476/SPP2177 and a grant (01GI10925) from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.). The analysis was conducted with data from the GNC (www.nako.de). The GNC is funded by the Federal Ministry of Education and Research (BMBF) (project funding reference numbers: 01ER1301A/B/C and 01ER1511D), federal states, and the Helmholtz Association with additional financial support by the participating universities and the institutes of the Leibniz Association. **Author contributions:** Conceptualization: T.H., F.S., C.L.S., J.N., K.H.M.-H., F.B., and J.M. Methodology: T.H. and J.M. Investigation: T.H. Visualization: T.H. Supervision: F.S., F.B., and J.M. Writing—original draft: T.H., F.B., and J.M. Writing—review and editing: All authors. Acquisition, curation, quality check, and provision of GNC data: H.-U.K., T.Ni., T.P., K.-H.J., L.U., A.P., S.R., T.K., N.H., H.V., L.K., S.N.W. **Competing interests:** N.H. is a stockholder of Siemens Healthineers. F.B. receives unrestricted research grants from Siemens Healthineers and speakers bureau from Siemens Healthineers. The authors declare that they have no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. In addition, the code of nnU-Net (*36*) used in this study is licensed under Apache License 2.0 and publicly available under https://github.com/MIC-DKFZ/nnUNet. The GNC data can be provided by the GNC data transfer site based on a written data request, pending scientific review, and writer cooperation agreement. Requests should be submitted to the GNC data transfer site (https://transfer.nako.de). The trained nnU-Net model parameters based on the GNC data used in this study are available under https://zenodo.org/record/7229667.

Submitted 8 June 2022
Accepted 11 April 2023
Published 12 May 2023
10.1126/sciadv.add0433

2.2 Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow

Haueise T, Stefan N, Schulz TJ, Schick F, Birkenfeld AL, Machann J.
Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow.
Z Med Phys. Published online January 30, 2023.

Methodological background and approach

In general, there are two different approaches for measuring PDFF in the bone marrow of the vertebral bodies: Averaging the signal intensity in circular ROIs, usually placed in a single slice in the center of the vertebral body under investigation, or averaging over the entire volume of the vertebral body. Principally, circular ROIs can be placed in the image more quickly than manual delineation of the vertebral bodies.

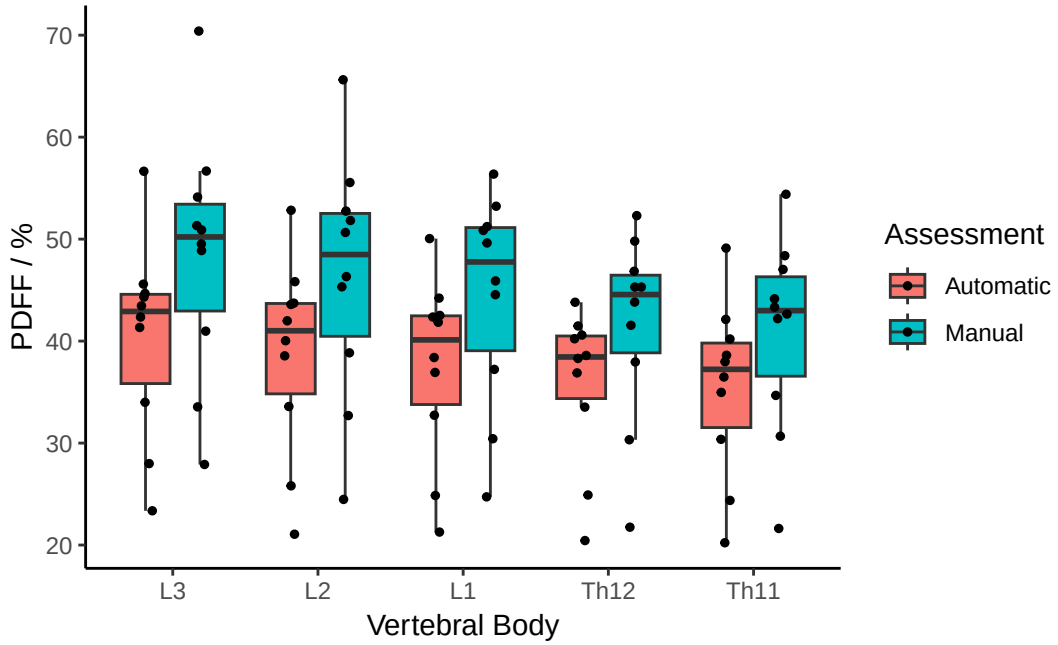

Figure 2.4: **Comparison between manual and automatic assessment.** Manual assessment of proton density fat fraction in vertebral bone marrow from six-point Dixon MRI using circular regions of interest (ROI) in a single slice in the center of vertebral bodies from the lumbar and thoracic spine (L3 to Th11) from 10 participants. ROI-based approach leads to systematic overestimation compared to volumetric assessment.

However, this approach assumes a homogeneous distribution of the BMAT in the vertebral body and neglects possible inhomogeneities or (intra- and interindividual) differences in its distribution. Three-dimensional imaging with a sufficiently small partition thickness makes it possible to detect these inhomogeneities.

In addition to interindividual differences, the vertebral bodies also differ intraindividually in terms of their volume or position in relation to axial images along the spine. This can lead to further inaccuracies in the quantification of BMAT.

Considering volumetric quantification, delineation of vertebral bodies should only include trabecular bone. Therefore, postprocessing, for example, binary erosion with a small structuring element, is employed to ensure that no compact bone or vertebral endplates are included.

Figure 2.4 illustrates the difference between manual ROI-based quantification of PDFF in BMAT for five different vertebral bodies (L3 to Th11) compared to the volumetric approach based on six-point Dixon MRI from 10 study participants. Details on imaging and study population are given in the publication. In general, the manual approach overestimated PDFF. Differences between the two methods ranged from 8.1%pt in L3 to 5.5%pt in Th11.

Using dependent *t*-tests with $\alpha = 0.05$ and Bonferroni correction for multiple comparisons, all found differences were significant (L3: $t_9 = 9.42$; $P < .001$, L2: $t_9 = 9.36$; $P < .001$, L1: $t_9 = 10.4$; $P < .001$, Th12: $t_9 = 5.46$; $P = .002$, Th11: $t_9 = 8.90$; $P < .001$).

Similar to sec. 2.1, a nnU-Net model was trained on 30 manually annotated data sets to automatically segment the vertebral bodies along the spine from fat-selective Dixon MRI. Before evaluating the segmentations, erosion with a 3x3 matrix-of-ones used as structuring element is applied.

Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow ☆

Tobias Haueise [a,b,c], **Norbert Stefan** [b,c,d], **Tim J. Schulz** [c,e,f], **Fritz Schick** [a,b,c], **Andreas L. Birkenfeld** [b,c,d], **Jürgen Machann** [a,b,c,*]

[a] Section on Experimental Radiology, Department of Diagnostic and Interventional Radiology, University Hospital Tübingen, Tübingen, Germany
[b] Institute for Diabetes Research and Metabolic Diseases, Helmholtz Munich at the University of Tübingen, Tübingen, Germany
[c] German Center for Diabetes Research (DZD), Tübingen, Germany
[d] Department of Diabetology, Endocrinology and Nephrology, University Hospital Tübingen, Tübingen, Germany
[e] Department of Adipocyte Development and Nutrition, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany
[f] Institute of Nutritional Science, University of Potsdam, Potsdam, Germany

Received 8 December 2022; accepted 19 December 2022

Abstract

This work proposes a method for automatic standardized assessment of bone marrow volume and spatial distribution of the proton density fat fraction (PDFF) in vertebral bodies. Intra- and interindividual variability in size and shape of vertebral bodies is a challenge for comparable interindividual evaluation and monitoring of changes in the composition and distribution of bone marrow due to aging and/or intervention. Based on deep learning image segmentation, bone marrow PDFF of single vertebral bodies is mapped to a cylindrical template and corrected for the inclination with respect to the horizontal plane. The proposed technique was applied and tested in a cohort of 60 healthy (30 males, 30 females) individuals. Obtained bone marrow volumes and mean PDFF values are comparable to former manual and (semi-)automatic approaches. Moreover, the proposed method allows shape-independent characterization of the spatial PDFF distribution inside vertebral bodies.

Keywords: Bone marrow; CSE MRI; Deep learning; PDFF; Segmentation

1 Introduction

The composition and regional distribution of hematopoietic (red) bone marrow (BM) in the vertebral bodies is variable and subject to age- and gender-dependent changes in healthy individuals (healthy aging) [1–4]. Furthermore, significant changes in composition can occur due to metabolic diseases (e.g. diabetes type 2), osteoporosis or hematological diseases [5–9]. In addition, there are various therapeutic strategies such as systemic chemotherapy and local radiotherapy that change the composition of BM [10–13].

In order to study the properties of BM and to monitor changes without radiation exposure, quantitative magnetic resonance imaging (qMRI) and magnetic resonance

☆ Supported in part by a grant (01GI0925) from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.). T.J.S. receives funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project No. 427826188 / SFB 1444. Methodical tools were developed with support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project No. 428224476/SPP 2177.

* Corresponding author: PD Dr. sc. hum. Jürgen Machann, Section on Experimental Radiology, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany.
E-mail: juergen.machann@med.uni-tuebingen.de (J. Machann).

spectroscopy (MRS) have been applied [14–17]. In contrast to single voxel ^{1}H-MRS, which has been considered the reference standard for the analysis of composition of BM but is limited to localized focal regions of a few cm^3 and therefore not able to assess the spatial distribution of fat and water inside vertebral bodies, quantitative imaging techniques can be used to measure the bone marrow proton density fat fraction (BM_{PDFF}) in a three-dimensional volume covering the complete anatomical region of interest. As vertebral bone marrow is characterized by short T2* times due to the presence of iron containing paramagnetic hematopoietic cells [18] and trabecular bone [17], multi-echo Dixon techniques with consideration of the spectral fat signal components and correction of the signal decay of the measured components are currently considered the gold standard for volumetric determination of BM_{PDFF} (e.g. 6-point Dixon [19]).

Quantitative assessment of BM_{PDFF} can be performed using different strategies: Evaluation of regions of interest (ROIs) (e.g. central circular ROIs [1,5]), manual segmentation of the bone marrow cavity [20,21], or automatic segmentation using deep learning incorporating multiple sagittal image slices [22]. However, all those approaches do not adequately consider the different size and interindividual variability in shape of the vertebral bodies: the regional distribution and inhomogeneity of BM_{PDFF} inside the vertebral body are often neglected. A standardized, shape-independent method which is able to monitor features of PDFF distribution inside vertebral bodies and to compare changes in subgroups regarding anthropometric and clinical parameters (e.g. insulin sensitivity) is still missing.

In this work, a shape-independent method for the standardized analysis of BM_{PDFF} inside vertebral bodies is presented. Therefore, a three-dimensional deep learning (DL)-based vertebral body segmentation model is used to generate shape-independent BM_{PDFF} distribution maps of the lumbar spine that allow, besides the quantification of BM_{PDFF}, an interindividual analysis of sub-regional distribution patterns inside the vertebral bodies.

2 Materials and methods

2.1 Subjects

For testing the proposed method, imaging data of a study cohort of 60 healthy Caucasian volunteers without bone malignancies in their medical history (30 females) was analyzed. The subjects were split in two age groups for males and females using the median age of 50 years as threshold (15 females <50 years, 33.9 ± 9.6 years, 24.9 ± 7.3 kg/m^2; 15 females >50 years, 59.9 ± 6.2 years, 26.1 ± 6.3 kg/m^2; 15 males <50 years, 32.6 ± 8.3 years, 26.3 ± 4.9 kg/m^2; 15 males >50 years, 60.0 ± 6.5 years, 26.9 ± 3.7 kg/m^2). Differences in BMI between age groups were not significant for both genders ($p = 0.4067$ and $p = 0.3837$, respectively).

The study was approved by the local ethics committee and written informed consent was obtained from all subjects prior to participation.

2.2 MR examinations

MR examinations were performed on a 3 T whole-body scanner (Magnetom Vida, Siemens Healthcare, Erlangen, Germany). Subjects were positioned head first in supine position with a spine-array coil mounted on the patient table of the scanner. For homogeneous coverage of the body trunk, an 18-channel body-array coil was placed on the lower abdomen. A 3D volumetric interpolated breath-hold examination (VIBE) 6-point chemical-shift-encoding (CSE) Dixon sequence covering the lumbar spine was executed using the following parameters: matrix size 160x104, field-of-view 380x313 mm, in-plane voxel size 1.2 x 1.2 mm, section thickness 3 mm, TE = 1.09, 2.46, 3.69, 4.92, 6.15 and 7.38 ms, TR = 13 ms, flip angle 4°, encoding acceleration with Caipirinha, factor 2 in both, phase-encoding and slice encoding directions, bandwidth 1078 Hz/pixel, acquisition time TA = 17 s (breath-hold). The PDFF maps were generated inline on the console of the scanner applying the vendor's algorithm, correcting for microscopic magnetic field inhomogeneities (T2*) and applying a multi-peak fat model [23].

For acquisition of training data for the DL model, a similar T1-weighted VIBE two-point Dixon sequence with 1.4 x 1.4 mm in-plane voxel size, 3 mm section thickness, TE = 1.23 ms and TR = 4.36 ms was applied.

2.3 Automated evaluation of BM_{PDFF} using deep learning-based image segmentation

For automated segmentation of the vertebral bodies, a DL-based 3D U-Net model (nnU-Net, full resolution configuration) [24] was derived from an independent stratified (gender, age, BMI) MRI data set of 30 subjects from local ongoing studies as described above (see Sec. 2.2). The manual labeling process of lumbar vertebrae was performed by a doctoral student (T.H.) under the supervision of two experienced medical physicists (>25 years of experience, F.S., J.M.).

The model was trained out-of-the-box (i.e. without the manual tuning of any hyperparameters) using a five-fold cross-validation scheme over 1000 epochs, providing the mean (i.e. the output of an ensemble) of the five independent models as resulting segmentation as suggested in [24].

The model-generated segmentations were evaluated using metrics of the class-wise confusion matrix as well as by the volumetric error (spatial accuracy of the model) compared to

the manual ground truth segmentation from the training set as well as on a manually segmented stratified random sample of 12 subjects from the study cohort.

For further validation of the obtained BM_{PDFF} quantification from the trained DL model, an independent, open access reference database of lumbar BM_{PDFF} values (MyoSegmenTUM spine) [25] was used.

After application of the trained DL model on the four groups of the study cohort, volumetric characterization (mean, min, max and SD of all voxels) of BM_{PDFF} was derived from the automatic segmentations of the 6-point Dixon MR images by averaging over all voxels of the segmented vertebral body.

2.4 Creation of standardized BM_{PDFF} distribution maps

Based on the automatically obtained segmentations (see Fig. 1A–C), each vertebral body is projected onto a standardized cylindrical shape using slice-wise linear interpolation of the PDFF map. After this registration of each vertebra to a standard cylindrical shape, three perpendicular BM_{PDFF} standardized distribution maps corresponding to the mid-transverse, mid-sagittal and mid-coronal plane of each vertebra can be derived and used for further analysis (see Fig. 1D).

The cylindrical shape is defined in terms of sampling points used for interpolation. In the mid-transverse plane of each vertebra, an elliptical base area is constructed using 20 and 16 sampling points on the major and minor semi-axes, respectively. First, the PDFF map is interpolated row-wise along the direction of the horizontal axis. Second, column-wise interpolation in the direction of the dorsoventral axis is performed. According to the elliptical template shape, the rows are resampled to yield the PDFF distribution map. The mid-coronal and mid-sagittal planes are constructed using the center-of-mass of the segmented vertebral body. Analogously, row-wise interpolation along the horizontal or dorsoventral axis is followed by interpolation in head-feet direction. Along the longitudinal axis, 20 sampling points are used.

To correct for the vertebral body inclination in the transverse plane, the segmented 3D PDFF image volume is

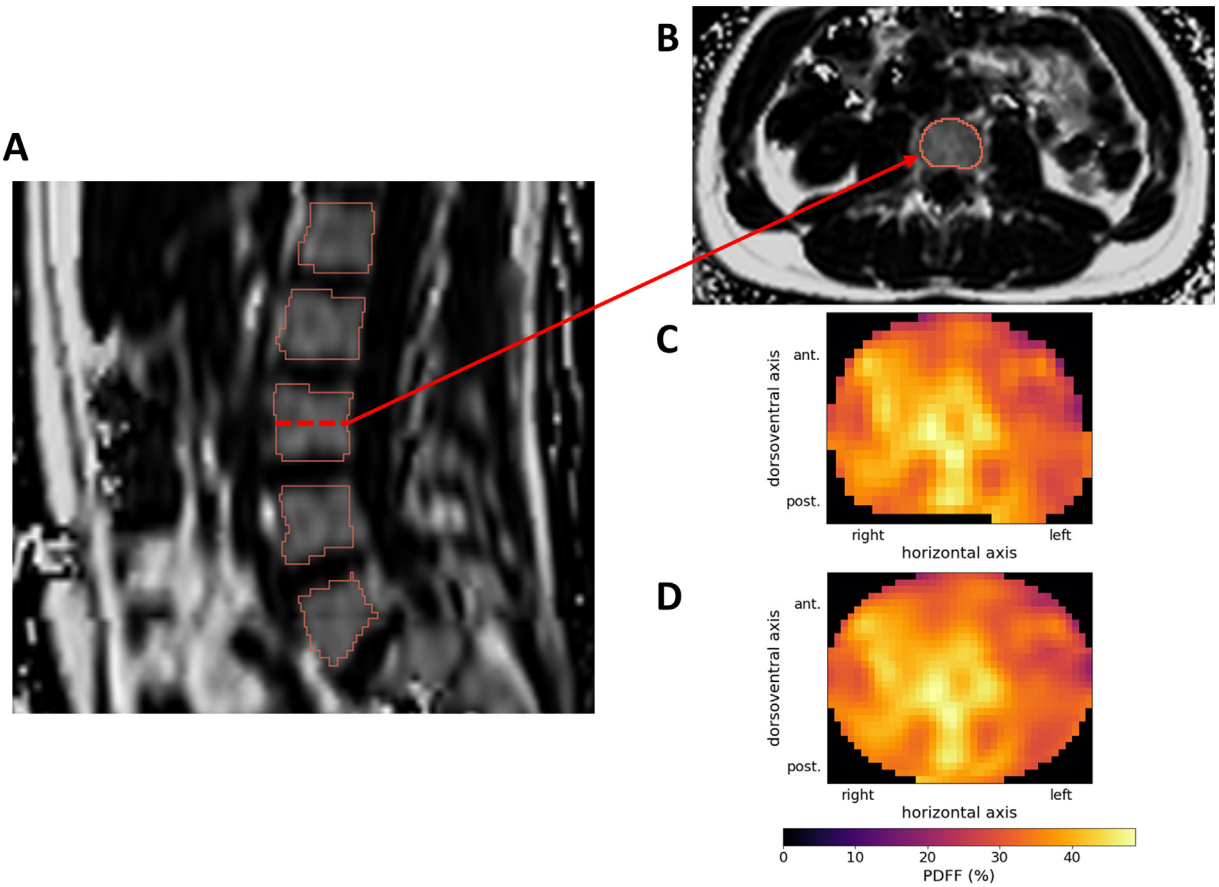

Figure 1. Workflow of the creation of standardized BM_{PDFF} distribution maps. Based on automated segmentation of vertebrae (**A**, **B**), raw BM_{PDFF} values (**C**) are mapped to a standardized cylindrical shape using linear interpolation along horizonal and dorsoventral axes (**D**).

rotated by detecting the inferior edge in the mid-sagittal plane of each vertebra. Its inclination angle is obtained using Hough transform of the segmentation map.

The extent to which the inclination correction has an influence is investigated by analyzing mean BM_{PDFF}, range and SD of BM_{PDFF}. SD is considered a measure of inhomogeneity of BM_{PDFF} inside the vertebral body.

The generated standardized distribution maps are compared to the corresponding raw PDFF image slices and the volumetric assessment from the segmentation model in terms of the metrics mentioned above.

2.5 Statistical tests

All data are reported as mean ± SD unless stated otherwise.

Performance of the segmentation model is evaluated using class-wise metrics of the confusion matrix, e.g. Dice similarity coefficient (DSC).

Bland-Altman plots are used to visualize the agreement between the automated volumetric quantification of BM_{PDFF} and the values derived from the proposed standardized distribution maps.

Welch's t-test or Mann-Whitney u-test are selected to compare the different strategies for quantification and assessment of differences in anthropometric measures, as appropriate. *p*-values <0.05 were considered statistically significant.

All statistical analyses were performed in Python 3.8 using SciPy 1.8.0.

3 Results

3.1 DL-based segmentation

The evaluation of the cross-validated training metrics shows that the segmentation model precisely segments the vertebral bodies of the lumbar spine, both in terms of mean DSC (0.894 ± 0.026) and mean volumetric error (1.41 ± 0.26 ml, corresponding to 8.63 ± 2.15 %) across all lumbar vertebrae. The cross-validation metrics are summarized in Table 1.

On a random sample of the study cohort (n = 12, see Sec. 2.3), mean DSC (0.901 ± 0.024) and mean volumetric error (1.52 ± 0.29 ml, corresponding to 7.48 ± 1.24 %) of all vertebrae of the lumbar spine were found comparable within the standard deviation.

External validation using MyoSegmenTUM spine reference data [25] shows good agreement in terms of BM_{PDFF} values. For female and male subjects <50 years, differences in sample size and anthropometrics are not significant (15 females <50 years, 29.9 ± 7.1 years, 26.0 ± 1.6 kg/m^2, p = 0.2039 and p = 0.5771, respectively; 15 males <50 years, 30.5 ± 4.9 years, 27.4 ± 2.8 kg/m^2, p = 0.4117 and p = 0.4901, respectively). In both groups, differences in BM_{PDFF} of the lumbar spine are also not significant from L5–L1 compared to the study cohort (p = 0.6840–0.8525 for females and p = 0.1031–0.8176 for males, respectively).

3.2 Standardized distribution maps

The correction for the inclination of the vertebral body did not result in any significant change of the obtained mean, range or SD of BM_{PDFF} compared to their initial value without the correction in both the generated distribution maps as well as in the corresponding PDFF image slices. One exception was found for L5 in the group of females >50 years (p = 0.0357 and p = 0.0293 for the generated maps and the PDFF image slices, respectively).

Mean BM_{PDFF} obtained from generated distribution maps do not show any significant differences across genders and age groups compared to the values obtained from the PDFF image slices and the volumetric assessment.

A significant overestimation of the minimum BM_{PDFF} (up to 6.1 % for females >50 years compared to the mean volumetric BM_{PDFF}) occurred in L1 for both genders and age groups (p = 0.0141 for females <50 years, p = 0.0017 for males <50 years, p < 0.001 for females >50 years, p = 0.0032 for males >50 years). For females in both age groups, a similar overestimation (up to 6.3 % for females >50 years) of the minimum BM_{PDFF} was found in L2. Correcting for the inclination angle reduces the difference to the volumetric minimum from 5.3 % for females >50 years up to non-significance in L5 for female subjects in both age groups (p = 0.1385 for females <50 years and p = 0.1664 for females >50 years, respectively).

The inhomogeneity of BM_{PDFF} as expressed by SD was significantly underestimated in L1 for all groups except females <50 years. Correcting for the inclination results in an approximation to the volumetric assessment but differences remain significant.

In some other cases (underestimation of SD in L2 and L3 for males >50 years, overestimation of the minimum in L4 for males <50 years), significant deviations from the corresponding volumetric assessment were found. By correction for the inclination, significant underestimations of SD in L4 for males of both age groups (p = 0.0817 for males <50 years, p = 0.0623 for males >50 years) in L5 for males <50 years (p = 0.4853) and of the maximum in L2 for males >50 years (p = 0.0639) are eliminated.

Except for the minimum BM_{PDFF}, all other metrics from the generated distribution maps are highly correlated (R^2 = 0,6813 for SD in males <50 years up to R^2 = 0,9812 for mean BM_{PDFF} in females <50 years) with the volumetric BM_{PDFF}. In most cases (except SD for females <50 years, maximum for females >50 years, mean and maximum for

Please cite this article as: T. Haueise, N. Stefan, T.J. Schulz et al., Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow, Z Med Phys, https://doi.org/10.1016/j.zemedi.2022.12.004

Table 1
Mean validation metrics of the class-wise confusion matrix and standard deviation of the five-fold cross-validated model training using a total of 30 annotated data sets.

	L5	L4	L3	L2	L1
Accuracy	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001
DSC	0.848 ± 0.190	0.916 ± 0.087	0.907 ± 0.076	0.898 ± 0.106	0.900 ± 0.080
FDR	0.148 ± 0.193	0.077 ± 0.103	0.100 ± 0.098	0.106 ± 0.125	0.103 ± 0.103
FNR	0.132 ± 0.121	0.087 ± 0.084	0.081 ± 0.079	0.092 ± 0.106	0.088 ± 0.010
FOR	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001
FPR	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001
Jaccard	0.767 ± 0.200	0.856 ± 0.133	0.837 ± 0.118	0.828 ± 0.144	0.826 ± 0.113
NPV	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001
Precision	0.852 ± 0.193	0.923 ± 0.103	0.900 ± 0.098	0.894 ± 0.125	0.897 ± 0.103
Recall	0.868 ± 0.122	0.913 ± 0.084	0.919 ± 0.079	0.908 ± 0.106	0.912 ± 0.010
TNR	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001	0.999 ± 0.001
Rel. error (%)	7.81 ± 7.62	6.07 ± 8.14	8.42 ± 7.27	8.85 ± 7.14	11.9 ± 7.79
Abs. error (ml)	1.65 ± 2.97	1.02 ± 1.31	1.33 ± 1.19	1.40 ± 1.18	1.65 ± 1.08

FDR false discovery rate, *FNR* false negative rate, *FOR* false omission rate, *FPR* false positive rate, *NPV* negative predictive value, *TNR* true negative rate.

males >50 years) the inclination correction leads to a slightly worsened correlation between the values obtained from the generated distribution maps and the respective volumetric assessment.

Regarding the mean BM_{PDFF}, Bland-Altman analysis reveals a low negative bias of the generated distribution maps compared to the volumetric assessment (see Fig. 2).

When applied to the study cohort, significant differences in mean BM_{PDFF} with respect to age are found in both genders and all vertebrae. Across genders, females and males <50 years show significant differences in mean BM_{PDFF} in all vertebrae except L4. Differences between females and males >50 years are not significant. Along the lumbar spine, mean BM_{PDFF} monotonically decreases from L5 to L1 (except in the group of females <50 years) (see Fig. 3).

The range of BM_{PDFF} extends from 35 % in L1 of females <50 years to 72 % in L5 of females >50 years. In females of both age groups, the range and inhomogeneity decrease from L5 to L1, whereas in males of both age groups the maximum range is found in L3 (see Table 2).

Qualitatively, the distribution maps show variable patterns across age groups and gender (see Fig. 4). The increase of mean BM_{PDFF} as well as the qualitative increase of inhomogeneity can be displayed for the purpose of visual inspection.

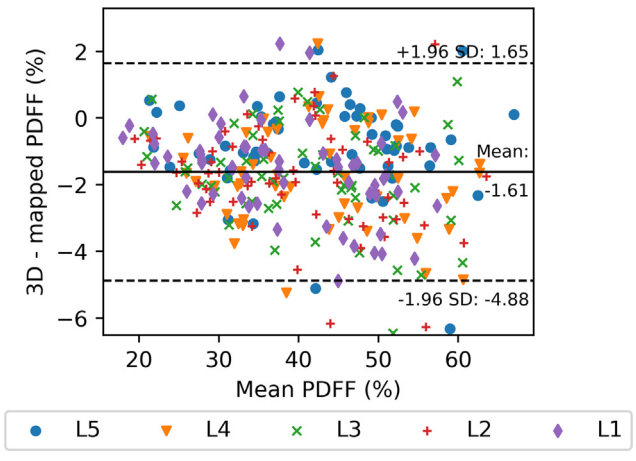

Figure 2. Bland-Altman plot showing agreement between BM_{PDFF} values derived from using all voxels from 3D segmentation and mean value of standardized distribution maps.

4 Discussion

Inhomogeneous composition and distribution of PDFF in vertebral (red) bone marrow (BM_{PDFF}) as well as intra- and interindividual spatial variability between vertebral bodies along the spine and across different subjects pose a particular challenge on the systematic, reproducible and standardized evaluation of BM_{PDFF}. In this work, the proposed method tackles these challenges by generating standardized distribution maps of vertebral bodies corrected for their inclination. These maps show high correlation with mean BM_{PDFF}, its range and its quantitative inhomogeneity obtained from the direct unstandardized volumetric assessment using DL-based image segmentation. The presented segmentation model allows for the precise and fully automated (and therefore time-saving) quantification of BM_{PDFF} similar to strategies reported in the literature [20,25–27]. The proposed standardized registration of vertebral bodies on a unified spatial structure allows both, direct shape-independent comparison of different vertebral bodies of an individual subject

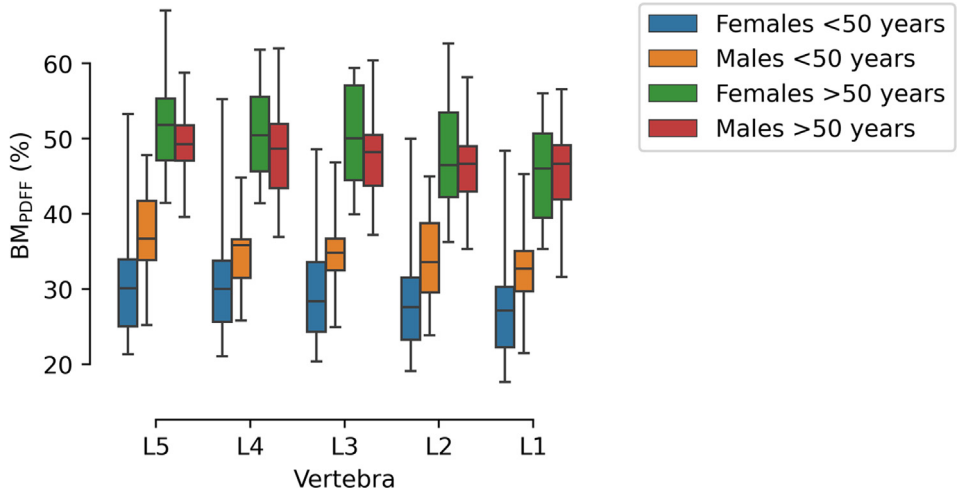

Figure 3. Quantified BM_{PDFF} of the lumbar spine in the four groups of the study cohort: females <50 years (blue), males <50 years (orange), females >50 years (green), males >50 years (red).

Table 2
BM_{PDFF} (%) in dependence of gender and age as obtained from the distribution maps.

		L5	L4	L3	L2	L1
Females <50	Min	9.9 ± 3.8	12.4 ± 5.9	9.9 ± 5.1	9.0 ± 3.9	7.6 ± 3.3
	Max	50.5 ± 14.6	50.6 ± 15.8	50.2 ± 16.5	46.0 ± 14.5	42.7 ± 11.4
	SD	6.5 ± 2.5	5.8 ± 2.2	6.4 ± 2.4	5.6 ± 2.4	4.9 ± 1.8
Males <50	Min	12.1 ± 5.2	12.8 ± 5.0	7.3 ± 4.9	10.1 ± 5.2	10.4 ± 5.3
	Max	57.4 ± 6.8	54.5 ± 6.4	60.9 ± 7.3	60.9 ± 14.1	54.8 ± 10.4
	SD	7.5 ± 1.4	6.8 ± 1.3	8.2 ± 1.1	7.6 ± 2.9	5.9 ± 1.2
Females >50	Min	18.0 ± 5.2	18.4 ± 6.0	17.0 ± 5.8	15.9 ± 6.1	13.3 ± 5.3
	Max	89.2 ± 15.7	85.6 ± 15.6	80.8 ± 13.7	78.4 ± 13.5	73.9 ± 9.4
	SD	11.3 ± 3.1	10.6 ± 3.0	10.5 ± 2.3	10.0 ± 2.5	9.2 ± 1.6
Males >50	Min	18.5 ± 8.4	16.6 ± 8.4	10.6 ± 5.8	11.4 ± 6.7	11.8 ± 4.2
	Max	82.0 ± 14.6	79.1 ± 12.0	76.3 ± 13.8	76.1 ± 7.7	71.6 ± 10.0
	SD	8.8 ± 3.2	9.1 ± 1.5	9.8 ± 1.6	9.8 ± 1.8	8.6 ± 1.4

as well as interindividual comparability or analyses in anthropometrically matched groups. As a downside, the method relies on DL-based image segmentation as an expensive (in terms of time and computational complexity) preprocessing step that also needs a manually segmented training data set. Training data were labeled in fat-selective images but – due to the heterogenous composition of bone cavity – delineation will probably lead to identical results using other contrasts as water selective, in-phase and/or opposed phase images available from the applied CSE sequence. However, a 3D dataset with sufficient spatial resolution is mandatory in order to avoid partial volume effects by inclusion of structures assigned to compact bone. Reported under- or overestimations of single metrics (e. g. minimum BM_{PDFF}) do not show a clear systematic, but are related to each other: overestimation of the minimum (as found in L1) decreases the range of possible values of BM_{PDFF} thereby reducing the standard deviation which is expressed in an underestimation of BM_{PDFF} inhomogeneity.

Spatial standardization to provide an interindividual comparison of MRI-assessed biomarkers is common practice in other body regions as CSE-based MRI enables reliable assessment and depiction of fat and, especially, its distribution in different organs. For example, the body structure is standardized for the assessment of whole-body adipose tissue topography along the longitudinal axis of the body [28]. When multi-point Dixon techniques are applied to organs such as the pancreas [29,30] or skeletal muscle [31,32] to quantify ectopic lipids, or for the quantification of intrahepatic lipids [23,33], the distribution of fat has mainly been taken under consideration in the pancreas differentiating fat content for head, body and tail [34] or skeletal muscle [32]. On the other hand, the liver is mostly considered as a whole by selecting a dedicated ROI in a

Please cite this article as: T. Haueise, N. Stefan, T.J. Schulz et al., Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow, Z Med Phys, https://doi.org/10.1016/j.zemedi.2022.12.004

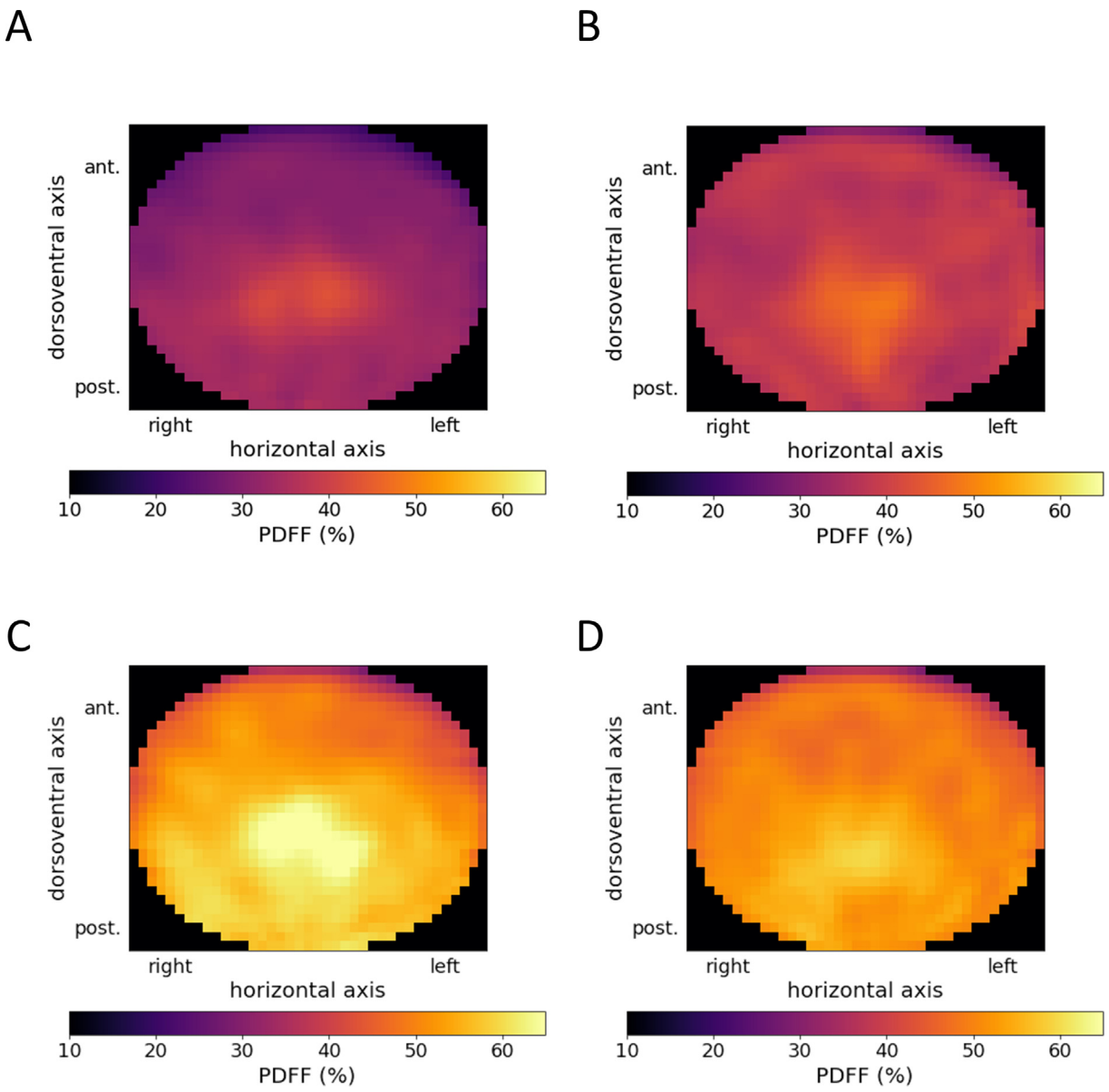

Figure 4. Standardized horizontal distribution maps of L3 showing qualitative differences between the four groups of the study cohort: females <50 years (**A**), males <50 years (**B**), females >50 years (**C**), males >50 years (**D**).

representative part of the tissue. Thus, both approaches are found in the evaluation of BM_{PDFF} [1–3,17].

Given the high reproducibility of CSE-based MRI between platforms of different field strengths and vendors [35,36], it furthermore enables detection of short-term [34] and long-term changes [37] in PDFF. Using the proposed method, the standardized analysis of short- and long-term changes might also be of interest for the clinical and epidemiological characterization of bone marrow, e.g. in the field of hematological diseases [12,13,38], osteoporosis [8,39,40], Modic change classification, i.e. pathological alterations by endplate degeneration with subchondral bone marrow changes [41] as well as after clinically indicated physical inactivity [42] or in analysis of interindividual and matched distribution patterns in large epidemiological studies like UK Biobank [43] or the German National Cohort [44].

4.1 Limitations

First, the relatively small number of subjects (n = 15, each with 5 vertebral bodies) in analyzed subgroups harms the generalizability of the results such that a potential broader variability including special cases may not be adequately covered. An epidemiological study studying the

Please cite this article as: T. Haueise, N. Stefan, T.J. Schulz et al., Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow, Z Med Phys, https://doi.org/10.1016/j.zemedi.2022.12.004

number and type of vertebral deformities [39] indicates that vertebral deformities occur mostly along the longitudinal axis (crush, wedge or biconcave deformities). Therefore, the proposed method based on the mid-transverse plane of the vertebrae should not be impaired. Furthermore, the study was not carried out to assess differences in age- and gender subgroups with high statistical significance, but to test a strategy for anatomically standardized assessment of BM_{PDFF} and its distribution in vertebral bodies. An extension to subjects with, e.g., osteoporosis is subject to future studies. Second, the study is based on the hypothesis that the deep learning-based automatic segmentation of vertebral bodies allow for an accurate quantification of BM_{PDFF}. Although this assumption seems reasonable, further research may focus on its verification. Third, there are no data on metabolic status and/or on bone health available in the study cohort which might allow an evaluation of spatial BM_{PDFF} distribution in relation to metabolic alterations (as insulin resistance, type 2 diabetes) and/or osteoporosis. Forth, the application of the method is only shown on the vertebrae of the lumbar spine. However, the translation to images covering the whole spine is possible.

5 Conclusion

The proposed method based on automated deep learning-based image segmentation can not only be used for robust and precise measurement of BM_{PDFF} based on the segmentation of the whole vertebral body but also to visualize its spatial distribution pattern inside the vertebral body. The method can serve as a tool to overcome biases in the assessment of representative BM_{PDFF} originating from varying quantification techniques and enables the systematic analysis of distribution patterns in groups of individuals. Further investigations of these distributions have the potential to deepen the understanding of the functionality of vertebral BM_{PDFF} and may thus help in evaluation of large-scale epidemiological studies as well as in follow-up analyses in patients with hematological diseases or therapy response.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank all study participants for their contribution and the members of Siemens Healthineers for their continuous support.

References

[1] Baum T, Yap SP, Dieckmeyer M, Ruschke S, Eggers H, Kooijman H, Rummeny EJ, Bauer JS, Karampinos DC. Assessment of whole spine vertebral bone marrow fat using chemical shift-encoding based water-fat MRI. J Magn Reson Imaging JMRI 2015;42:1018–1023.

[2] Aoki T, Yamaguchi S, Kinoshita S, Hayashida Y, Korogi Y. Quantification of bone marrow fat content using iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): reproducibility, site variation and correlation with age and menopause. Br J Radiol 2016;89:20150538.

[3] Griffith JF, Yeung DKW, Ma HT, Leung JCS, Kwok TCY, Leung PC. Bone marrow fat content in the elderly: a reversal of sex difference seen in younger subjects. J Magn Reson Imaging JMRI 2012;36:225–230.

[4] Ruschke S, Syväri J, Dieckmeyer M, Junker D, Makowski MR, Baum T, Karampinos DC. Physiological variation of the vertebral bone marrow water T2 relaxation time. NMR Biomed 2021;34:e4439.

[5] Ma Q, Cheng X, Hou X, Yang Z, Ma D, Wang Z. Bone Marrow Fat Measured by a Chemical Shift-Encoded Sequence (IDEAL-IQ) in Patients With and Without Metabolic Syndrome. J Magn Reson Imaging JMRI 2021;54:146–153.

[6] Bredella MA, Gill CM, Gerweck AV, Landa MG, Kumar V, Daley SM, Torriani M, Miller KK. Ectopic and serum lipid levels are positively associated with bone marrow fat in obesity. Radiology 2013;269:534–541.

[7] Li G, Xu Z, Gu H, Li X, Yuan W, Chang S, Fan J, Calimente H, Hu J. Comparison of chemical shift-encoded water-fat MRI and MR spectroscopy in quantification of marrow fat in postmenopausal females. J Magn Reson Imaging JMRI 2017;45:66–73.

[8] Gassert FT, Kufner A, Gassert FG, Leonhardt Y, Kronthaler S, Schwaiger BJ, Boehm C, Makowski MR, Kirschke JS, Baum T, Karampinos DC, Gersing AS. MR-based proton density fat fraction (PDFF) of the vertebral bone marrow differentiates between patients with and without osteoporotic vertebral fractures. Osteoporos Int J Establ Result Coop Eur Found Osteoporos Natl Osteoporos Found USA 2022;33:487–496.

[9] Machann J, Stefan N, Schick F. (1)H MR spectroscopy of skeletal muscle, liver and bone marrow. Eur J Radiol 2008;67:275–284.

[10] Bolan PJ, Arentsen L, Sueblinvong T, Zhang Y, Moeller S, Carter JS, Downs LS, Ghebre R, Yee D, Froelich J, Hui S. Water-fat MRI for assessing changes in bone marrow composition due to radiation and chemotherapy in gynecologic cancer patients. J Magn Reson Imaging JMRI 2013;38:1578–1584.

[11] Carmona R, Pritz J, Bydder M, Gulaya S, Zhu H, Williamson CW, Welch CS, Vaida F, Bydder G, Mell LK. Fat composition changes in bone marrow during chemotherapy and radiation therapy. Int J Radiat Oncol Biol Phys 2014;90:155–163.

[12] Schick F, Einsele H, Weiß B, Forster J, Lutz O, Kanz L, Claussen CD. Assessment of the composition of bone marrow prior to and following autologous BMT and PBSCT by magnetic resonance. Ann Hematol 1996;72:361–370.

[13] Pereira PL, Schick F, Einsele H, Farnsworth CT, Kollmansberger C, Mattke A, Duda SH, Claussen CD. MR tomography of the bone marrow changes after high-dosage chemotherapy and autologous peripheral stem-cell transplantation. ROFO Fortschr Geb Rontgenstr Nuklearmed 1999;170:251–257.

[14] Griffith JF, Yeung DKW, Antonio GE, Lee FKH, Hong AWL, Wong SYS, Lau EMC, Leung PC. Vertebral bone mineral density, marrow perfusion, and fat content in healthy men and men with osteoporosis: dynamic contrast-enhanced MR imaging and MR spectroscopy. Radiology 2005;236:945–951.

[15] Schick F, Bongers H, Jung WI, Skalej M, Lutz O, Claussen CD. Volume-selective proton MRS in vertebral bodies. Magn Reson Med 1992;26:207–217.

[16] Li X, Kuo D, Schafer AL, Porzig A, Link TM, Black D, Schwartz AV. Quantification of vertebral bone marrow fat content using 3 Tesla MR spectroscopy: reproducibility, vertebral variation, and applications in osteoporosis. J Magn Reson Imaging JMRI 2011;33:974–979.

[17] Karampinos DC, Melkus G, Baum T, Bauer JS, Rummeny EJ, Krug R. Bone marrow fat quantification in the presence of trabecular bone: initial comparison between water-fat imaging and single-voxel MRS. Magn Reson Med 2014;71:1158–1165.

[18] Le Ster C, Gambarota G, Lasbleiz J, Guillin R, Decaux O, Saint-Jalmes H. Breath-hold MR measurements of fat fraction, T1, and T2 * of water and fat in vertebral bone marrow. J Magn Reson Imaging JMRI 2016;44:549–555.

[19] Karampinos DC, Ruschke S, Dieckmeyer M, Eggers H, Kooijman H, Rummeny EJ, Bauer JS, Baum T. Modeling of T2* decay in vertebral bone marrow fat quantification. NMR Biomed 2015;28:1535–1542.

[20] Sollmann N, Dieckmeyer M, Schlaeger S, Rohrmeier A, Syvaeri J, Diefenbach MN, Weidlich D, Ruschke S, Klupp E, Franz D, Rummeny EJ, Zimmer C, Kirschke JS, Karampinos DC, Baum T. Associations Between Lumbar Vertebral Bone Marrow and Paraspinal Muscle Fat Compositions-An Investigation by Chemical Shift Encoding-Based Water-Fat MRI. Front Endocrinol 2018;9:563.

[21] Schmeel FC, Vomweg T, Träber F, Gerhards A, Enkirch SJ, Faron A, Sprinkart AM, Schmeel LC, Luetkens JA, Thomas D, Kukuk GM. Proton density fat fraction MRI of vertebral bone marrow: Accuracy, repeatability, and reproducibility among readers, field strengths, and imaging platforms. J Magn Reson Imaging JMRI 2019;50:1762–1772.

[22] Zhou J, Damasceno PF, Chachad R, Cheung JR, Ballatori A, Lotz JC, Lazar AA, Link TM, Fields AJ, Krug R. Automatic Vertebral Body Segmentation Based on Deep Learning of Dixon Images for Bone Marrow Fat Fraction Quantification. Front Endocrinol 2020;11:612.

[23] Zhong X, Nickel MD, Kannengiesser SAR, Dale BM, Kiefer B, Bashir MR. Liver fat quantification using a multi-step adaptive fitting approach with multi-echo GRE imaging. Magn Reson Med 2014;72:1353–1365.

[24] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18:203–211.

[25] Burian E, Rohrmeier A, Schlaeger S, Dieckmeyer M, Diefenbach MN, Syväri J, Klupp E, Weidlich D, Zimmer C, Rummeny EJ, Karampinos DC, Kirschke JS, Baum T. Lumbar muscle and vertebral bodies segmentation of chemical shift encoding-based water-fat MRI: the reference database MyoSegmenTUM spine. BMC Musculoskelet Disord 2019;20:152.

[26] Baum T, Rohrmeier A, Syväri J, Diefenbach MN, Franz D, Dieckmeyer M, Scharr A, Hauner H, Ruschke S, Kirschke JS, Karampinos DC. Anatomical Variation of Age-Related Changes in Vertebral Bone Marrow Composition Using Chemical Shift Encoding-Based Water-Fat Magnetic Resonance Imaging. Front Endocrinol 2018;9:141.

[27] Dieckmeyer M, Junker D, Ruschke S, Mookiah MRK, Subburaj K, Burian E, Sollmann N, Kirschke JS, Karampinos DC, Baum T. Vertebral Bone Marrow Heterogeneity Using Texture Analysis of Chemical Shift Encoding-Based MRI: Variations in Age, Sex, and Anatomical Location. Front Endocrinol 2020;11:555931.

[28] Machann J, Thamer C, Schnoedt B, Haap M, Haring H-U, Claussen CD, Stumvoll M, Fritsche A, Schick F. Standardized assessment of whole body adipose tissue topography by MRI. J Magn Reson Imaging JMRI 2005;21:455–462.

[29] Tushuizen ME, Bunck MC, Pouwels PJ, Bontemps S, van Waesberghe JHT, Schindhelm RK, Mari A, Heine RJ, Diamant M. Pancreatic fat content and beta-cell function in men with and without type 2 diabetes. Diabetes Care 2007;30:2916–2921.

[30] Heni M, Machann J, Staiger H, Schwenzer NF, Peter A, Schick F, Claussen CD, Stefan N, Häring H-U, Fritsche A. Pancreatic fat is negatively associated with insulin secretion in individuals with impaired fasting glucose and/or impaired glucose tolerance: a nuclear magnetic resonance study. Diabetes Metab Res Rev 2010;26:200–205.

[31] Patzelt L, Junker D, Syväri J, Burian E, Wu M, Prokopchuk O, Nitsche U, Makowski MR, Braren RF, Herzig S, Diaz MB, Karampinos DC. MRI-Determined Psoas Muscle Fat Infiltration Correlates with Severity of Weight Loss during Cancer Cachexia. Cancers 2021;13:4433.

[32] Kiefer LS, Fabian J, Rospleszcz S, Lorbeer R, Machann J, Kraus MS, Roemer F, Rathmann W, Meisinger C, Heier M, Nikolaou K, Peters A, Storz C, Diallo TD, Schlett CL, Bamberg F. Distribution patterns of intramyocellular and extramyocellular fat by magnetic resonance imaging in subjects with diabetes, prediabetes and normoglycaemic controls. Diabetes Obes Metab 2021;23:1868–1878.

[33] Park CC, Hooker C, Hooker JC, Bass E, Haufe W, Schlein A, Covarrubias Y, Heba E, Bydder M, Wolfson T, Gamst A, Loomba R, Schwimmer J, Hernando D, Reeder SB, Middleton M, Sirlin CB, Hamilton G. Assessment of a high-SNR chemical-shift-encoded MRI with complex reconstruction for proton density fat fraction (PDFF) estimation overall and in the low-fat range. J Magn Reson Imaging JMRI 2019;49:229–238.

[34] Machann J, Hasenbalg M, Dienes J, Wagner R, Sandforth A, Fritz V, Birkenfeld AL, Nikolaou K, Kullmann S, Schick F, Heni M. Short-Term Variability of Proton Density Fat Fraction in Pancreas and Liver Assessed by Multiecho Chemical-Shift Encoding-Based MRI at 3 T. J Magn Reson Imaging JMRI 2022. https://doi.org/10.1002/jmri.28084.

[35] Wu B, Han W, Li Z, Zhao Y, Ge M, Guo X, Wu X. Reproducibility of Intra- and Inter-scanner Measurements of Liver Fat Using Complex Confounder-corrected Chemical Shift Encoded MRI at 3.0 Tesla. Sci Rep 2016:6:19339.

[36] Hu HH, Yokoo T, Bashir MR, Sirlin CB, Hernando D, Malyarenko D, Chenevert TL, Smith MA, Serai SD, Middleton MS, Henderson WC, Hamilton G, Shaffer J, Shu Y, Tkach JA, Trout AT, Obuchowski N, Brittain JH, Jackson EF, Reeder SB, Quantitative Imaging Biomarkers Alliance RSNA, Biomarker Committee PDFF. Linearity and Bias of Proton Density Fat Fraction as a Quantitative Imaging Biomarker: A Multicenter, Multiplatform, Multivendor Phantom Study. Radiology 2021;298:640–651.

[37] Lavynenko O, Abdul-Ghani M, Alatrach M, Puckett C, Adams J, Abdelgani S, Alkhouri N, Triplitt C, Clarke GD, Vasquez JA, Li J, Cersosimo E, Gastaldelli A, DeFronzo RA. Combination therapy with pioglitazone/exenatide/metformin reduces the prevalence of hepatic fibrosis and steatosis: The efficacy and durability of initial combination therapy for type 2 diabetes (EDICT). Diabetes Obes Metab 2022;24:899–907.

[38] Machann J, Pereira PL, Einsele H, Kanz L, Claussen CD, Schick F. The MR characterization of the composition of the hematopoietic bone marrow. The findings in generalized neoplasms and the monitoring of therapy. Radiol 2000;40:700–709.

[39] Ismail AA, Cooper C, Felsenberg D, Varlow J, Kanis JA, Silman AJ, O'Neill TW, the European Vertebral Osteoporosis Study Group. Number and Type of Vertebral Deformities: Epidemiological Characteristics and Relation to Back Pain and Height Loss. Osteoporos Int 1999;9:206–213.

[40] Sollmann N, Kirschke JS, Kronthaler S, Boehm C, Dieckmeyer M, Vogele D, Kloth C, Lisson CG, Carballido-Gamio J, Link TM, Karampinos DC, Karupppasamy S, Beer M, Krug R, Baum T. Imaging of the Osteoporotic Spine - Quantitative Approaches in Diagnostics and for the Prediction of the Individual Fracture Risk. ROFO Fortschr Geb Rontgenstr Nuklearmed 2022;194:1088–1099.

[41] Modic MT, Pflanze W, Feiglin DH, Belhobek G. Magnetic resonance imaging of musculoskeletal infections. Radiol Clin North Am 1986;24:247–258.

[42] Liu T, Melkus G, Ramsay T, Sheikh A, Laneuville O, Trudel G. Bone Marrow Reconversion With Reambulation: A Prospective Clinical Trial. Invest Radiol 2021;56:215–223.

[43] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779.

[44] German National Cohort (GNC) Consortium. The German National Cohort: aims, study design and organization. Eur J Epidemiol 2014;29:371–382.

2.3 Comparison of the accuracy of commercial two-point and multi-echo Dixon MRI for quantification of fat in liver, paravertebral muscles, and vertebral bone marrow

Haueise T, Schick F, Stefan N, Machann J.
Comparison of the accuracy of commercial two-point and multi-echo Dixon MRI for quantification of fat in liver, paravertebral muscles, and vertebral bone marrow. *Eur J Radiol*. 2024;172:111359.

Methodological background and approach

While the previous sections focused on the tissue level of body composition (**wangFivelevelModelNew1992**), the perspective in the following section shifts more to the molecular level. Dixon-type imaging is used for quantification of fat fractions (FF) in tissues. However, besides other sources of bias outlined in sec.

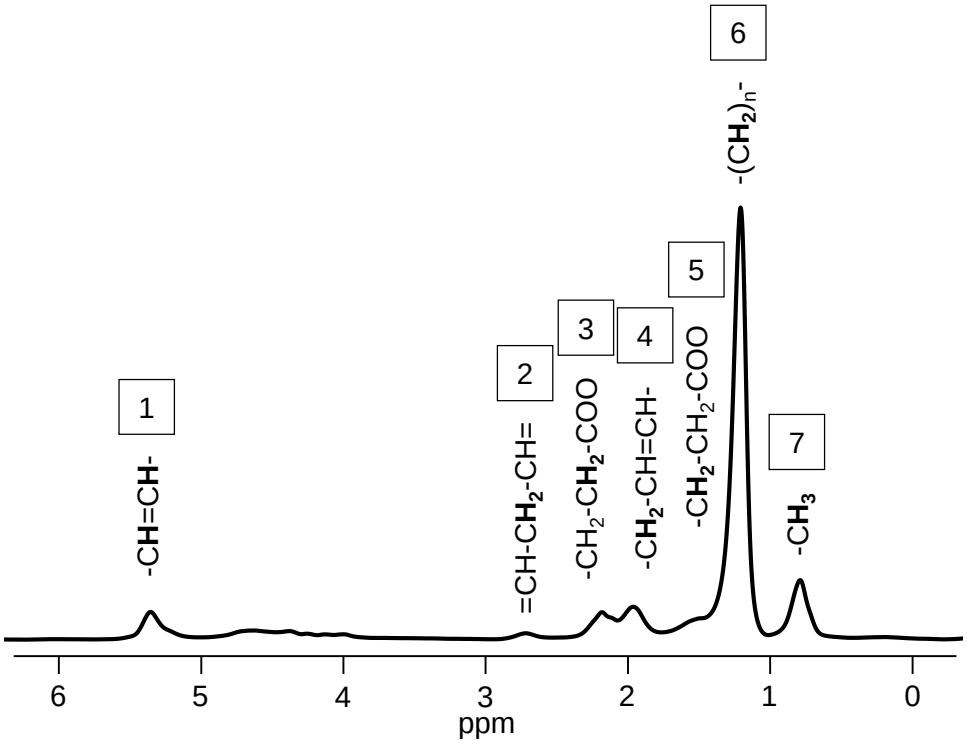

Figure 2.5: **Spectrum from visceral adipose tissue.** Relevant metabolites for characterization of triglyceride composition are shown in a spectrum from visceral adipose tissue from a 38-year old woman. The dominant methylene peak is numbered 6. Other numbers mark vinyl (1), diallylic (2), α-methylene (3), allylic (4), β-methylene (5), and methyl (7) peaks.

1.1 that can be addressed during image acquisition, for example, T1 bias by using small flip angles, the underlying signal model is a key element in chemical shift-encoded imaging that can lead to substantial bias in FF quantification (Hernando et al. 2010; Reeder et al. 2009). Selection of the signal model involves three key decisions: 1) magnitude or complex fitting, 2) single or multi-peak modeling of triglyceride spectrum, 3) correction for effective transverse relaxation (T2*) decay (Hernando et al. 2010).

Two-point Dixon methods usually use a simple signal model considering fat as a single peak at -420 Hz relative to water at 3 T (corresponding to the methylene peak labeled 6 in Figure 2.5) (Eggers et al. 2014) without incorporating T2* effects (Bydder et al. 2008). In fact, the spectrum of triglycerides is more complex and it is possible to resolve up to 10 different peaks depending on field strength and T2* (Peterson et al. 2021). Hence, using a multi-peak signal model, for example, including seven peaks as the method used in the following does (Zhong et al. 2014), demonstrated agreement with gold-standard MRS, not only in terms of correlation but also regarding slope (Reeder et al. 2009). Figure 2.5 shows an example of a typical spectrum from VAT showing relevant metabolites for characterization of triglyceride composition. T2* decay leads to a decrease of the amplitudes of water and fat components in the MR signal with increasing echo time leading to nonnegligible errors in FF estimation (Bydder et al. 2008) that can be of clinical importance, for example, in the quantification of liver fat in the presence of iron (Westphalen et al. 2007).

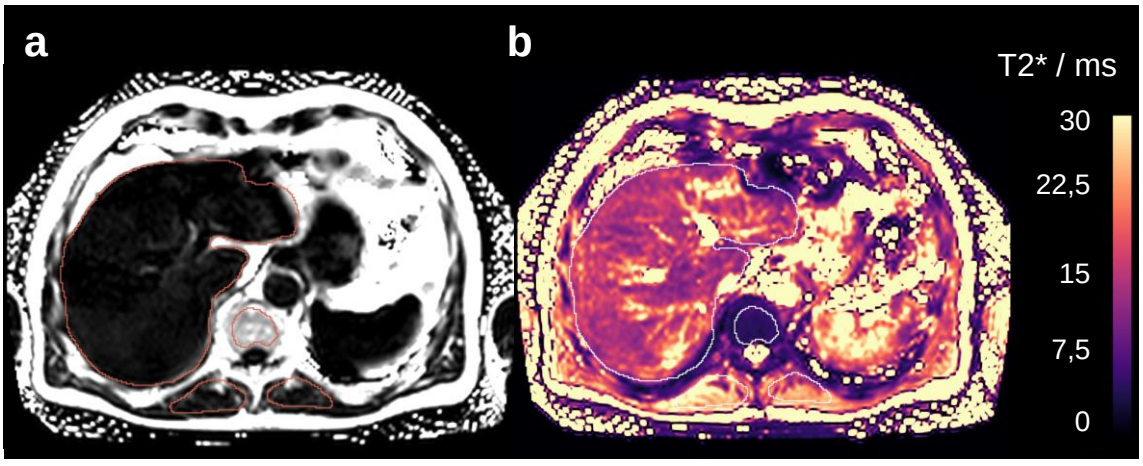

Figure 2.6: **Proton density fat fraction and T2*.** PDFF (a) and T2* map (b) from a 64 year-old man highlighting the liver, Th10 vertebra, and a pair of skeletal muscles. Mean PDFF of the liver was 5.4%, 7.6% in skeletal muscle, and 35.4% in Th10. Mean T2* was 20.3 ms in the liver, 11.4 ms in skeletal muscle, and 5.5 ms in Th10.

When applying commercial methods, the signal model's implementation is a closed system, and potential bias from its design choices, characterization of methods like two-point Dixon imaging is necessary prior to application. For example, considering quantification of BMAT in the NAKO, only T1-weighted two-point Dixon imaging is available for the complete spine leading to confounded FF estimates. Since different tissues can have different T2* times, the following characterization of a commercial two-point Dixon technique also includes tissues with shorter T2*, i.e., skeletal muscles or BMAT. Figure 2.6 shows PDFF (a) and T2* map (b) highlighting the liver, vertebral body Th10, and a pair of skeletal muscles.

European Journal of Radiology 172 (2024) 111359

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Comparison of the accuracy of commercial two-point and multi-echo Dixon MRI for quantification of fat in liver, paravertebral muscles, and vertebral bone marrow

Tobias Haueise [a,b], Fritz Schick [a,b,c], Norbert Stefan [b,c,d], Jürgen Machann [a,b,c,*]

[a] Institute for Diabetes Research and Metabolic Diseases, Helmholtz Munich at the University of Tübingen, Tübingen, Germany
[b] German Center for Diabetes Research (DZD), Tübingen, Germany
[c] Section on Experimental Radiology, Department of Diagnostic and Interventional Radiology, University Hospital Tübingen, Tübingen, Germany
[d] Department of Diabetology, Endocrinology and Nephrology, University Hospital Tübingen, Tübingen, Germany

ARTICLE INFO

Keywords:
Magnetic resonance imaging
Dixon MRI
Vertebral bone marrow
Fat fraction estimation

ABSTRACT

Purpose: Excess fat accumulation contributes significantly to metabolic dysfunction and diseases. This study aims to systematically compare the accuracy of commercially available Dixon techniques for quantification of fat fraction in liver, skeletal musculature, and vertebral bone marrow (BM) of healthy individuals, investigating biases and sex-specific influences.

Method: 100 healthy White individuals (50 women) underwent abdominal MRI using two-point and multi-echo Dixon sequences. Fat fraction (FF), proton density fat fraction (PDFF) and T2* values were calculated for liver, paravertebral muscles (PVM) and vertebral BM (Th8–L5). Agreement and systematic deviations were assessed using linear correlation and Bland-Altman plots.

Results: High correlations between FF and PDFF were observed in liver ($r = 0.98$ for women; $r = 0.96$ for men), PVM ($r = 0.92$ for women; $r = 0.93$ for men) and BM ($r = 0.97$ for women; $r = 0.95$ for men). Relative deviations between FF and PDFF in liver (18.92 % for women; 13.32 % for men) and PVM (1.96 % for women; 11.62 % for men) were not significant. Relative deviations in BM were significant (38.13 % for women; 27.62 % for men). Bias correction using linear models reduced discrepancies. T2* times were significantly shorter in BM (8.72 ms for women; 7.26 ms for men) compared to PVM (13.45 ms for women; 13.62 ms for men) and liver (29.47 ms for women; 26.35 ms for men).

Conclusion: While no significant differences were observed for liver and PVM, systematic errors in BM FF estimation using two-point Dixon imaging were observed. These discrepancies – mainly resulting from organ-specific T2* times – have to be considered when applying two-point Dixon approaches for assessment of fat content. As suitable correction tools, linear models could provide added value in large-scale epidemiological cohort studies. Sex-specific differences in T2* should be considered.

1. Introduction

Excess accumulation of ectopic fat within various anatomical compartments in the human body has been recognized as a significant indicator of metabolic dysfunction, insulin resistance, and various diseases, such as type-2 diabetes, atherosclerosis, multiple myeloma or osteoporosis [1–3]. Non-invasive and reliable quantification of fat accumulation within specific tissues, such as bone marrow (BM) in vertebral bodies, musculature, and liver, is pivotal for understanding their impact on health and disease [4–7].

Magnetic resonance imaging (MRI) and volume localized spectroscopy (MRS) have emerged as important tools in this endeavor. Modern MRI techniques provide non-invasive assessment of macroscopic fat distribution [8,9] as well as ectopic fat accumulation within organs and tissues [10–12] with high spatial resolution. Among the various MRI methods available, Dixon-type imaging has gained substantial attention for its ability to distinguish between water and fat content within tissues [13] even in regions with inhomogeneous magnetic field distribution. Further, Dixon MRI can be used to measure fat fraction (FF) at the voxel level.

* Corresponding author at: Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany.
E-mail address: juergen.machann@med.uni-tuebingen.de (J. Machann).

https://doi.org/10.1016/j.ejrad.2024.111359
Received 1 February 2024; Accepted 3 February 2024
Available online 5 February 2024

However, the choice of Dixon MRI sequence and the methods employed for postprocessing of recorded imaging data can significantly affect the accuracy and reliability of FF measurements as reflected, e. g., in the compromise between acquisition time, image resolution, and accuracy.

In contrast to two-point Dixon imaging, which acquires images at two echo times (in-phase/opposed-phase) enabling calculation of fat-selective images, but requires an internal reference for (magnitude-based) FF estimation, multi-echo Dixon imaging allows for absolute quantification of fat fraction defined as proton density fat fraction (PDFF). Commercially available multi-echo Dixon sequences have been optimized for PDFF quantification in the liver by implementation of T1-insensitive acquisition, T2*-correction, and consideration of the spectral complexity of fat [14]. For this purpose, they are validated in terms of linearity, precision and reproducibility in multisite, multivendor phantom studies [15–17] and *in vivo* [18–20]. As reviewed by Starekova et al., widespread application of imaging-based PDFF assessment justify the assumption of this as a reference [20].

PDFF quantification is also of interest for other organs and tissues, e. g. paravertebral muscles (PVM) or vertebral bone marrow (BM) [21,22]. However, lipids in red BM in vertebral bodies are metabolically distinct from other fat depots [23] and their assessment is confounded, as paramagnetic hematopoietic cells and the presence of trabecular structures influence the microscopic magnetic field distribution inside the vertebral bodies [24]. Additionally, water and lipid signals in red BM often show similar spectral signal intensities [25]. BM fat is subject of several studies analyzing its associations, e. g., with body composition, osteoporosis or degeneration of intervertebral discs [26–28]. Therefore, precise quantification of vertebral BM fat content using Dixon-based MR sequences is necessary.

Aim of the present study is the systematic comparison of two-point Dixon imaging with multi-echo Dixon imaging as a reference in a cohort of healthy individuals to investigate and assess biases of fat quantification in liver, PVM, and BM of thoracic and lumbar vertebral bodies.

2. Materials and methods

In this analysis, data from 100 healthy volunteers who self-reported as White individuals (50 women, sex defined based on self-report) participating in ongoing studies involving metabolic imaging, which were related to the Tübingen Diabetes Family Study [6], in Tübingen, Germany, were included. All individual studies were approved by the Ethics Committee of the University Hospital Tübingen and written informed consent was obtained from all participants prior to participation.

MR examinations were performed on a 3 T whole-body scanner (Magnetom Vida, Siemens Healthcare, Erlangen, Germany). Subjects were positioned head first in supine position on a 24-channel table-integrated spine-array coil. For homogeneous coverage of the body trunk, two 18-channel body-array coils were placed on chest and lower abdomen. A 3D volumetric interpolated breath-hold examination (VIBE) two-point Dixon sequence (regular product functionality) and a multi-echo Dixon VIBE sequence (LiverLab option) using six echoes were applied in three axial slabs covering the trunk. Sequence parameters of both acquisitions are summarized in Table 1.

Two-point FF maps from two-point measurement were calculated offline using fat-selective (F) and water-selective (W) images by voxel-wise application of the formula $FF = F/(F + W)$. Multi-echo PDFF maps were generated inline on the console of the scanner by the vendor's algorithm, correcting for microscopic magnetic field inhomogeneities by modeling of an effective transverse relaxation time (T2*) and spectrally complex fat dephasing [14].

For objective measurements without manual and subjective placement of single ROIs, fat quantification was performed using automatically generated segmentation masks of liver [29], ten vertebral bodies of the thoracic and lumbar spine (Th8–L5) [30], and PVM (right/left erector spinae and psoas major muscles) [31] by averaging voxel fat fraction map values inside each segmented fat compartment. By applying two-dimensional erosion in the axial plane using a 3x3 matrix as structuring element, edge pixels were excluded. T2* was measured from the corresponding maps using the mean value over the same generated segmentation masks.

Table 1
Sequence parameters. MRI acquisition parameters as applied on a whole-body scanner (Magnetom Vida, Siemens Healthcare, Erlangen, Germany).

	Two-point Dixon	Multi-echo Dixon
Matrix size	320x161	160x104
Field of view / mm x mm	550x395	380x313
In-plane resolution / mm x mm	1.4 x 1.4	1.2 x 1.2
Slice thickness / mm	3	3
Partitions per slab	80	80
TE / ms	1.23, 2.46	1.09, 2.46, 3.69, 4.92, 6.15, 7.38
TR / ms	4.36	13
Flip angle / Degree	9	4
Bandwidth / Hz/Pixel	1042	1078
TA / s	12.6	16.2

TE: Echo times TR: Repetition time TA: Acquisition time.

All data are reported as mean (SD) unless stated otherwise. Linear correlation analysis and Bland-Altman plots were used to quantify and visualize the bias and agreement between FF and PDFF. Differences are expressed as percentages ([FF-PDFF]/mean %). Deviations from FF compared to PDFF are reported as "relative" percentages ([FF-PDFF]/PDFF %) unless stated otherwise. Welch's *t*-test was selected to test for statistically significant differences between women and men. Holm-Bonferroni method was applied to correct p-values for multiple testing when comparing vertebral bodies. $p < 0.05$ was considered statistically significant. All statistical analyses were performed in Python 3.8 using SciPy 1.8.0 and statsmodels 0.13.2.

3. Results

Imaging errors (e. g. partial fat–water swaps in the liver) and errors during computation of PDFF led to the exclusion of five participants. One participant was excluded due to a potentially pathologically low PDFF in BM, not fulfilling the assumption of a "healthy volunteer". The analyzed study population is characterized in Table 2. An exemplary coronal PDFF map is shown in Fig. 1.

FF estimation showed high correlation with PDFF in the liver ($r = 0.98$, for women; $r = 0.96$, for men, see Fig. 2a), in PVM ($r = 0.92$, for women; $r = 0.93$, for men, see Fig. 2b) and BM averaged along the spine (including vertebral bodies Th8–L5; $r = 0.97$, for women; $r = 0.95$, for men, see Fig. 2c). Correlation strength was significantly different for men and women ($p < 0.05$). In the liver, fitted sex-specific linear

Table 2
Study population. Anthropometric data of the study population. Range of the values is presented in square brackets.

	Women	Men
N	47	47
Age / years	42.7 (14.8) [20–67]	46.1 (15.9) [23–76]
Height /cm	167.9 (7.1) [148.5–185.0]	180.9 (8.2)*** [163.0–198.2]
Weight / kg	75.7 (19.1) [46.2–112.0]	87.5 (16.0)** [59.0–144.6]
BMI / kg/m^2	27.1 (7.5) [15.5–40.4]	26.8 (4.6) [18.9–37.9]

Sex differences: $^{**}p < 0.01$, $^{***}p < 0.001$.

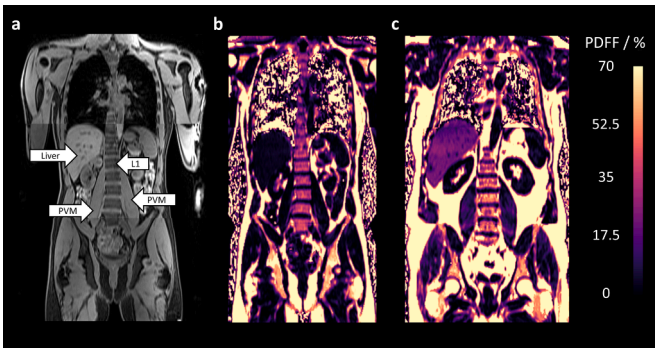

Fig. 1. Exemplary imaging data. Water-selective image slice from two-point acquisition (**a**) highlighting analyzed regions-of-interest and corresponding PDFF map (**b**) of a 28-year old man (BMI 22.3 kg/m^2, 3.50 % liver PDFF, 3.71 % muscle PDFF, 33.90 % mean bone marrow PDFF). (**c**) Coronal PDFF map of a 58-year old man with high liver and muscle PDFF (BMI 35.2 kg/m^2, 20.4 % liver PDFF, 8.64 % muscle PDFF, 39.27 % mean bone marrow PDFF).

regression models predicting PDFF from FF along with 95 % CI of regression parameters were: PDFF = 0.83 [0.79 to 0.88] x FF + 0.15 [−0.17 to 0.48] for women; PDFF = 0.83 [0.76 to 0.90] x FF + 0.61 [−0.07 to 1.28] for men. In PVM, the fitted models were: PDFF = 1.38 [1.20 to 1.56] x FF − 2.34 [−3.49 to −1.20] for women; PDFF = 1.29 [1.14 to 1.44] x FF − 2.23 [−3.20 to −1.26] for men. In BM, the fitted models were: PDFF = 0.86 [0.85 to 0.88] x FF − 5.89 [−6.82 to −4.97] for women; PDFF = 0.81 [0.79 to 0.84] x FF − 1.19 [−2.32 to −0.05] for men.

As detailed in Table 3, differences between FF and PDFF in liver and PVM are not significant for both sexes ($p = 0.41$ and $p = 0.79$ for women; $p = 0.68$ and $p = 0.25$ for men, respectively). Liver FF was increased compared to corresponding PDFF by 18.92 % in women and by 13.32 % in men. In PVM, FF was increased by 1.96 % in women and by 11.62 % in men. In contrast, FF in vertebral BM was significantly increased compared to PDFF by an average over all analyzed vertebral bodies of 38.13 % in women (ranging from 34.98 % in L5 to 41.00 % in Th9) and by an average of 27.62 % in men (ranging from 25.65 % in Th9 to 29.73 % in Th12). Across both sexes, bias in liver FF was 12.21 %, 5.27 % in PVM and 27.78 % in BM (see Fig. 3). With increasing mean fat content, bias in PVM and BM decreased (see Fig. 3b and 3c). In PVM, this even led to a shift from overestimation to underestimation of PDFF (see Fig. 3b). For mean fat fractions above 5 %, bias in the liver was almost constant (see Fig. 3a).

Considering FF estimation alone, apparently significant differences were found between different vertebral bodies, which could not be confirmed in PDFF quantification. In women and men, significant differences between PDFF in lower lumbar vertebrae (L4 and L5) and the mid thoracic spine (Th8–Th10) were found. In men, L1 and L2 were also significantly different from Th8–Th10 (see Fig. 4).

As measured from multi-echo Dixon, differences in T2* times between different vertebral bodies were not significant for women and men. Mean T2* in BM and liver were significantly higher for women compared to men (8.72 ms for women, 7.26 ms for men, $p < 0.001$, in BM; 29.47 ms for women, 26.35 ms for men, $p < 0.05$, in the liver; see Table 4). Sex differences in PVM were not significant ($p = 0.74$). For both sexes, mean T2* in BM was significantly lower compared to PVM and liver ($p < 0.001$).

Derived sex-specific linear regression models for BM PDFF (see Fig. 2c) can be used to correct for the bias in FF estimation. After application, FF overestimation was reduced to 0.04 % along the spine and across both sexes (see Fig. 5b). Additionally, apparently significant differences between FF and PDFF in the vertebral bodies, cannot be removed by linear correction.

4. Discussion

This study systematically compared two-point Dixon and a multi-echo Dixon sequences for the quantification of fat fraction in liver, paravertebral muscles (PVM) and bone marrow (BM) in vertebral bodies of the thoracic and lumbar spine (Th8–L5) in healthy subjects.

Table 3
Measured fat fractions. Fat fraction measurements (in "absolute" %) in liver, paravertebral muscles and vertebral bone marrow (Th8–L5) using a two-point and multi-echo Dixon sequence.

	Women		Men	
In %	PDFF	FF	PDFF	FF
Liver	4.73 (4.00)	5.48 (4.72)	6.29 (5.96)	6.84 (6.91)
PVM	6.55 (1.93)	6.46 (1.30)	5.90 (1.83)	6.28 (1.32)
Th8	30.51 (10.61)	41.92 (12.02)$^{+++}$	33.62 (7.79)	42.49 (9.69)$^{+++}$
Th9	29.66 (9.24)	41.02 (10.87)$^{+++}$	33.79 (7.80)	42.30 (9.63)$^{+++}$
Th10	31.03 (9.99)	42.10 (11.66)$^{+++}$	34.42 (7.64)	43.26 (9.98)$^{+++}$
Th11	32.05 (10.08)	43.71 (12.52)$^{+++}$	36.10 (8.09)	45.84 (9.72)$^{+++}$
Th12	34.49 (11.65)	47.02 (12.81)$^{+++}$	38.06 (8.44)	48.95 (9.39)$^{+++}$
L1	35.20 (10.99)	48.23 (12.25)$^{+++}$	38.88 (8.45)	49.81 (9.24)$^{+++}$
L2	36.39 (11.24)	49.08 (12.16)$^{+++}$	40.18 (8.40)	51.19 (9.22)$^{+++}$
L3	38.11 (11.66)	50.91 (12.56)$^{+++}$	41.22 (8.62)	52.18 (9.27)$^{+++}$
L4	39.54 (12.33)	52.32 (13.19)$^{+++}$	41.35 (8.71)	52.78 (9.88)$^{+++}$
L5	40.39 (13.18)	53.46 (14.28)$^{+++}$	43.54 (8.94)	54.61 (9.65)$^{+++}$

Intra-sex sequence differences: $^{+++}p < 0.001$; PVM: paravertebral muscles, PDFF: proton density fat fraction, FF: fat fraction from two-point Dixon

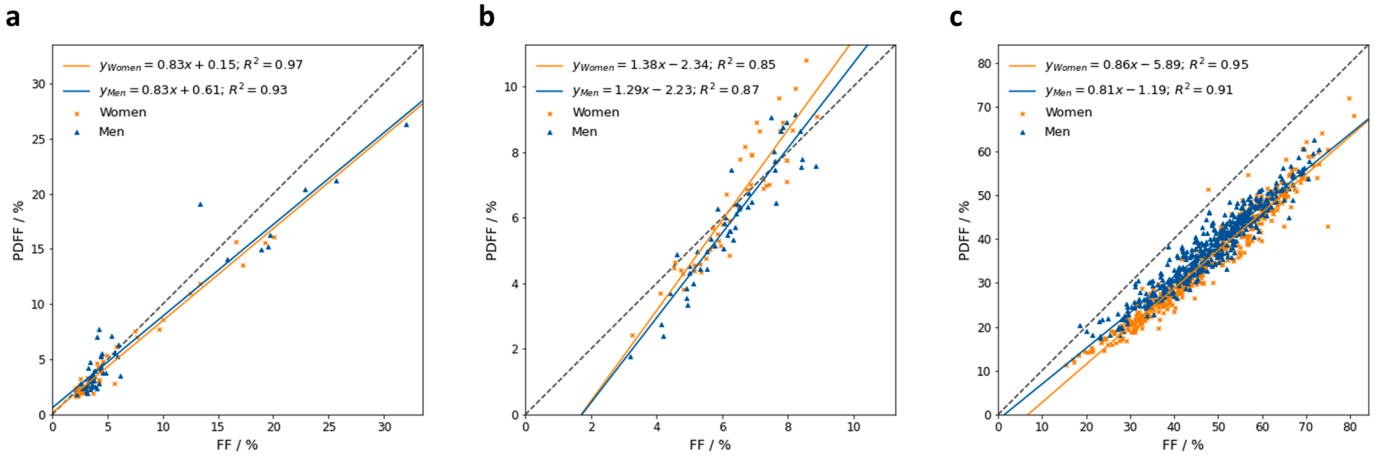

Fig. 2. Comparison of FF and PDFF. Comparison of two-point FF estimation with multi-echo PDFF quantification in the liver (**a**), paravertebral muscles (**b**) and in vertebral bone marrow (Th8–L5) (**c**).

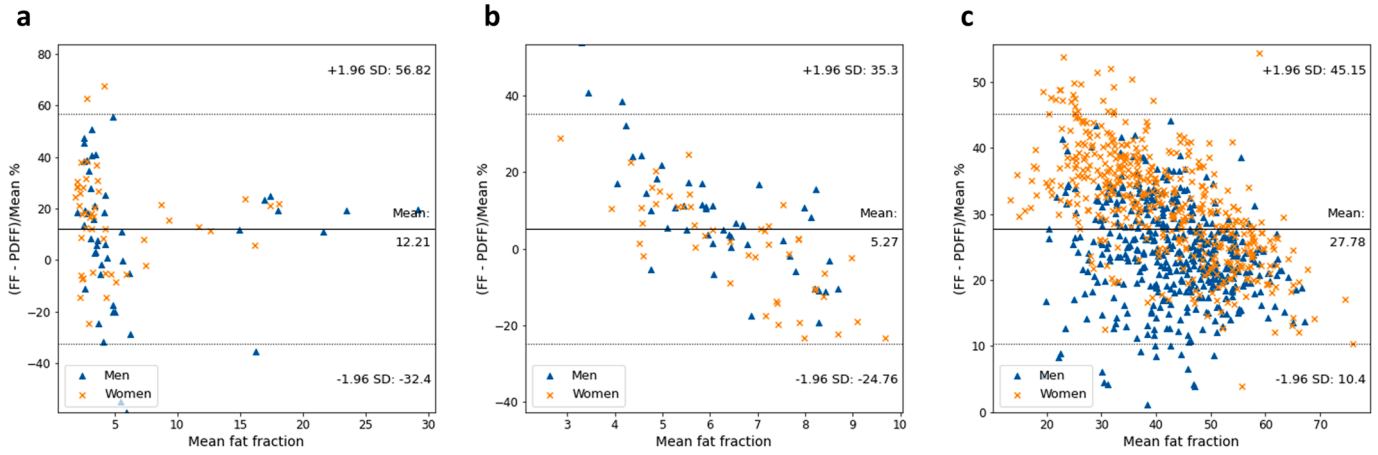

Fig. 3. Agreement of FF and PDFF. Bland-Altman plot comparing relative deviations between two-point FF estimation with multi-echo PDFF quantification in the liver (**a**), paravertebral muscles (**b**) and in vertebral bone marrow (Th8–L5) (**c**).

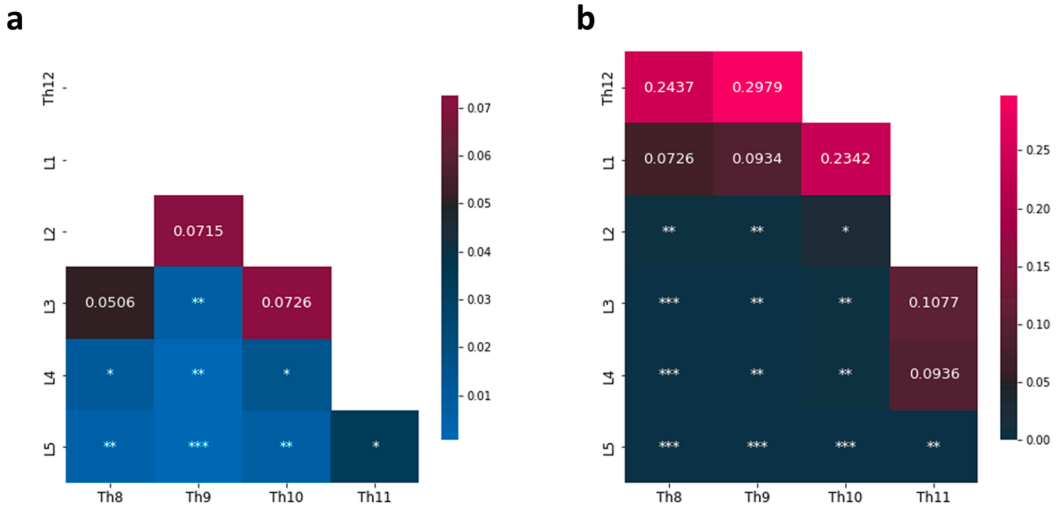

Fig. 4. Apparently significant differences. Holm-Bonferroni corrected *p*-values of apparently significant differences between FF and PDFF in vertebral bodies ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$) in women (**a**) and men (**b**).

Table 4
T2*. Mean T2* times for analyzed fat compartments as obtained from automatic calculation during image reconstruction using multi-echo Dixon sequence.

In ms	Women	Men
Liver	29.47 (5.22)	26.35 (8.51)*
PVM	13.45 (2.38)	13.62 (2.43)
Th8	8.99 (1.52)	7.59 (1.50)***
Th9	8.95 (1.47)	7.58 (1.49)***
Th10	8.69 (1.30)	7.11 (1.47)***
Th11	8.46 (1.24)	7.17 (1.47)***
Th12	8.78 (1.33)	7.26 (1.45)***
L1	8.76 (1.12)	7.39 (1.38)***
L2	8.59 (1.01)	7.25 (1.39)***
L3	8.56 (1.00)	7.12 (1.27)***
L4	8.50 (0.99)	7.11 (1.23)***
L5	8.90 (2.44)	7.02 (1.20)***

Sex differences: $^{*}p < 0.05$, $^{***}p < 0.001$; PVM: paravertebral muscles.

Significant differences in the estimation of vertebral BM fat fraction, despite very high correlation, were found. Using sex-specific linear regression models (as shown in Fig. 2c) for the correction of two-point FF in BM, bias between the corrected FF values and the multi-echo PDFF values can be removed (see Fig. 5). However, apparently significant differences between FF of single vertebral bodies could not be removed in the present analysis. This suggests that multi-echo measurements seem to be more suitable for in-depth analysis of BM fat beyond mere fat quantification.

Deviations between "confounder-corrected" fat quantification based on multi-echo Dixon PDFF maps and simple estimations from two echoes can stem from multiple sources: First, differences in T1 for the water and fat components of different tissues. As reviewed by Bojorquez et al., at 3 T, T1 in the liver ranges from 725 to 809 ms, lies between 898 and 1420 ms in muscle and is with a range between 106 and 586 ms notably shorter in BM [32]. More specifically, de Bazelaire et al. reported 809 ms for liver, 898 ms for paravertebral muscles and 586 ms for bone marrow in L4 using an inversion-recovery method and multiple spin-echo technique [33]). Second, different fat-to-water ratios can lead to under- or overestimation [34], and will scale with the flip angle of the acquisition, as well as with flip angle heterogeneity, especially in vertebral bodies due to trabecular structures. Although the applied two-point sequence is characterized by stronger T1-weighting compared to the multi-echo sequence, a resulting T1-effect is assumed to be less pronounced in BM due to shorter T1 (see above). Third, when not accounted for, the spectral complexity of fat will also cause the deviations to be dependent on the choice of echo times. T2* relaxation will reduce the signal intensity of later echoes, so the order of in-phase and opposed-phase echoes will become important. Especially inside the vertebral bodies,

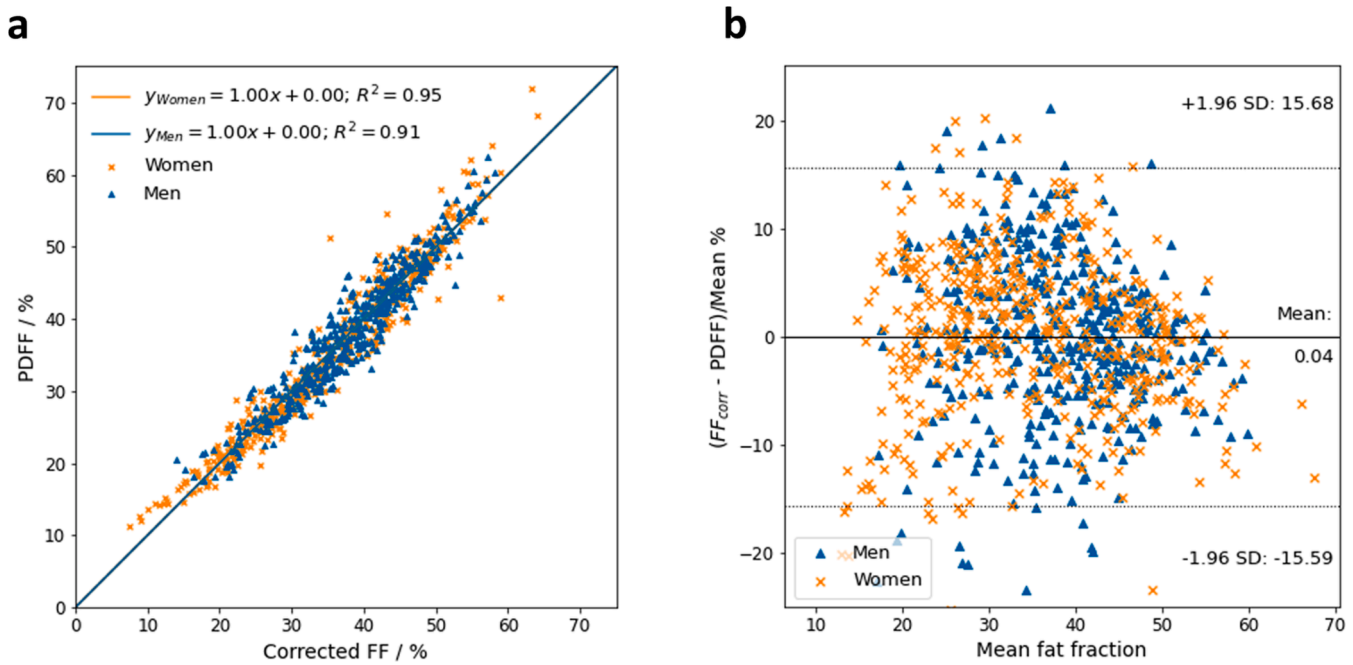

Fig. 5. Comparison and agreement of corrected FF. Comparison (**a**) and Bland-Altman plot (**b**) of corrected two-point FF and multi-echo PDFF in vertebral bone marrow (Th8–L5) as obtained from sex-specific linear regression models y = 0.86x-5.89 for women and y = 0.81x-1.19 for men.

heterogeneous trabecular structures and paramagnetic hematopoietic cells directly affect T2*.

This study has some limitations. First, only two different MR sequences from one manufacturer are considered. However, as images are taken from ongoing studies and the sequences used are also applied in large epidemiological cohort studies like KORA or *NAKO Gesundheitsstudie* [35,36], the results are of interest for a wide range of applications, as multiple studies investigated MRI-assessed BM fat and its correlation to body fat distribution, physical activity, as well as metabolic diseases, osteoporosis or its characterization after transplantation [3,26,28,37–40]. In other large-scale epidemiological cohort studies, such as UK Biobank, two-point Dixon imaging is also applied [41]. Furthermore, the underlying problem is in principle independent of sequence parameters and manufacturers: two-point measurements always have the problem of limited potential for confounder correction. However, no quantitative statement can be made about other manufacturers, but the results of the study can provide a qualitative indication of the expected level of error and the need for correction. Particularly in the liver, T2* can shorten dramatically in disease (iron overload) [42]. Thus, a fixed-factor T2* correction in two-point Dixon reconstruction may not be feasible. Second, the consideration of multi-echo Dixon PDFF as a baseline for comparison is not a true gold standard, as, inter alia, the adjustment for T2* correction is not ideal due to generally shorter effective relaxation times and owing to potential differences in the spectral composition of fat in bone marrow, which is calibrated for the liver [25]. Third, significant sex differences in FF inside the vertebral bodies, as well as in T2*, are not accounted for in two-point reconstruction. Finally, it has to be mentioned that the findings are limited to healthy subjects and might differ in patients with chronic liver disease [20], patients with neuromuscular disorders (muscular dystrophies) [21], or patients with hematological diseases undergoing cytostatic treatment or radiotherapy [40,43], where micro- and macroscopic alterations of the respective tissue have to be considered.

5. Conclusion

This study suggests that using two-point Dixon imaging for quantification of fat fraction in vertebral bone marrow reveals higher systematic errors compared to liver and/or muscle. Due to the high correlation between two-point and multi-echo FF estimation, simple tools such as linear models could be used to compensate for these systematic errors, as long as the acquisition parameters remain unchanged, when analyzed in large-scale epidemiological studies. Significant sex differences in T2* inside the vertebral bodies suggest the need for sex-specific reconstructions.

Funding

This work was supported in part by a grant of the German Research Foundation [428224476/SPP2177]; and in part by a grant from the Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.) [01GI0925].

CRediT authorship contribution statement

Tobias Haueise: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Fritz Schick:** Supervision, Writing – review & editing. **Norbert Stefan:** Resources, Writing – review & editing. **Jürgen Machann:** Conceptualization, Investigation, Methodology, Project administration, Resources, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank all participants who took part in the studies and the staff in this research program. Additionally, we thank Stephan Kannengießer from Siemens Healthineers for his valuable support.

References

[1] I.J. Neeland, R. Ross, J.-P. Després, et al., Visceral and ectopic fat, atherosclerosis, and cardiometabolic disease: a position statement, Lancet Diabetes Endocrinol. 7 (2019) 715–725, https://doi.org/10.1016/S2213-8587(19)30084-1.

[2] C. Falank, H. Fairfield, M.R. Reagan, Signaling interplay between bone marrow adipose tissue and multiple myeloma cells, Front. Endocrinol. 7 (2016) 67, https://doi.org/10.3389/fendo.2016.00067.

[3] K.M. Beekman, G. Duque, A. Corsi, et al., Osteoporosis and bone marrow adipose tissue, Curr. Osteoporos. Rep. 21 (2023) 45–55, https://doi.org/10.1007/s11914-022-00768-1.

[4] K.A. Britton, C.S. Fox, Ectopic fat depots and cardiovascular disease, Circulation 124 (2011), https://doi.org/10.1161/CIRCULATIONAHA.111.077602.

[5] N. Aaron, S. Costa, C.J. Rosen, L. Qiang, The implications of bone marrow adipose tissue on inflammaging, Front. Endocrinol. 13 (2022) 853765, https://doi.org/10.3389/fendo.2022.853765.

[6] N. Stefan, A. Fritsche, F. Schick, H.-U. Häring, Phenotypes of prediabetes and stratification of cardiometabolic risk, Lancet Diabetes Endocrinol. 4 (2016) 789–798, https://doi.org/10.1016/S2213-8587(16)00082-6.

[7] B.H. Goodpaster, B.C. Bergman, A.M. Brennan, L.M. Sparks, Intermuscular adipose tissue in metabolic disease, Nat. Rev. Endocrinol. 19 (2023) 285–298, https://doi.org/10.1038/s41574-022-00784-2.

[8] M. Borga, J. West, J.D. Bell, et al., Advanced body composition assessment: from body mass index to body composition profiling, J. Investig. Med. off Publ. Am. Fed. Clin. Res. 66 (2018) 1–9, https://doi.org/10.1136/jim-2018-000722.

[9] T. Haueise, F. Schick, N. Stefan, et al., Analysis of volume and topography of adipose tissue in the trunk: results of MRI of 11,141 participants in the German National Cohort, Sci. Adv. 9 (2023) eadd0433, https://doi.org/10.1126/sciadv.add0433.

[10] J. Machann, M. Hasenbalg, J. Dienes, et al., Short-term variability of proton density fat fraction in pancreas and liver assessed by multiecho chemical-shift encoding-based MRI at 3 T, J. Magn. Reson. Imaging JMRI (2022), https://doi.org/10.1002/jmri.28084.

[11] J. Machann, C. Thamer, B. Schnoedt, et al., Hepatic lipid accumulation in healthy subjects: a comparative study using spectral fat-selective MRI and volume-localized1H-MR spectroscopy, Magn. Reson. Med. 55 (2006) 913–917, https://doi.org/10.1002/mrm.20825.

[12] L.S. Kiefer, J. Fabian, S. Rospleszcz, et al., Assessment of the degree of abdominal myosteatosis by magnetic resonance imaging in subjects with diabetes, prediabetes and healthy controls from the general population, Eur. J. Radiol. 105 (2018) 261–268, https://doi.org/10.1016/j.ejrad.2018.06.023.

[13] H. Eggers, P. Börnert, Chemical shift encoding-based water-fat separation methods, J. Magn. Reson. Imaging JMRI 40 (2014) 251–268, https://doi.org/10.1002/jmri.24568.

[14] X. Zhong, M.D. Nickel, S.A.R. Kannengiesser, et al., Liver fat quantification using a multi-step adaptive fitting approach with multi-echo GRE imaging, Magn. Reson. Med. 72 (2014) 1353–1365, https://doi.org/10.1002/mrm.25054.

[15] H.H. Hu, T. Yokoo, M.R. Bashir, et al., Linearity and bias of proton density fat fraction as a quantitative imaging biomarker: a multicenter, multiplatform, multivendor phantom study, Radiology 298 (2021) 640–651, https://doi.org/10.1148/radiol.2021202912.

[16] T. Hayashi, K. Fukuzawa, H. Yamazaki, et al., Multicenter, multivendor phantom study to validate proton density fat fraction and T2* values calculated using vendor-provided 6-point DIXON methods, Clin. Imaging 51 (2018) 38–42, https://doi.org/10.1016/j.clinimag.2018.01.011.

[17] J.K. Jang, S.S. Lee, B. Kim, et al., Agreement and reproducibility of proton density fat fraction measurements using commercial MR sequences across different platforms: a multivendor, multi-institutional phantom experiment, Invest. Radiol. 54 (2019) 517–523, https://doi.org/10.1097/RLI.0000000000000561.

[18] T. Yokoo, S.D. Serai, A. Pirasteh, et al., Linearity, Bias, and precision of hepatic proton density fat fraction measurements by using MR imaging: a meta-analysis, Radiology 286 (2018) 486–498, https://doi.org/10.1148/radiol.2017170550.

[19] C. Zhan, S. Olsen, H.C. Zhang, et al., Detection of hepatic steatosis and iron content at 3 Tesla: comparison of two-point Dixon, quantitative multi-echo Dixon, and MR spectroscopy, Abdom. Radiol. NY 44 (2019) 3040–3048, https://doi.org/10.1007/s00261-019-02118-9.

[20] J. Starekova, D. Hernando, P.J. Pickhardt, S.B. Reeder, Quantification of liver fat content with CT and MRI: state of the art, Radiology 301 (2021) 250–262, https://doi.org/10.1148/radiol.2021204288.

[21] J. Burakiewicz, C.D.J. Sinclair, D. Fischer, et al., Quantifying fat replacement of muscle by quantitative MRI in muscular dystrophy, J. Neurol. 264 (2017) 2053–2067, https://doi.org/10.1007/s00415-017-8547-3.

[22] K. Engelke, O. Chaudry, L. Gast, et al., Magnetic resonance imaging techniques for the quantitative analysis of skeletal muscle: State of the art, J. Orthop. Transl. 42 (2023) 57–72, https://doi.org/10.1016/j.jot.2023.07.005.

[23] K.J. Suchacki, A.A.S. Tavares, D. Mattiucci, et al., Bone marrow adipose tissue is a unique adipose subtype with distinct roles in glucose homeostasis, Nat. Commun. 11 (2020) 3097, https://doi.org/10.1038/s41467-020-16878-2.

[24] C. Le Ster, G. Gambarota, J. Lasbleiz, et al., Breath-hold MR measurements of fat fraction, T1, and T2 * of water and fat in vertebral bone marrow, J. Magn. Reson. Imaging JMRI 44 (2016) 549–555, https://doi.org/10.1002/jmri.25205.

[25] J. Machann, N. Stefan, F. Schick, (1)H MR spectroscopy of skeletal muscle, liver and bone marrow, Eur. J. Radiol. 67 (2008) 275–284, https://doi.org/10.1016/j.ejrad.2008.02.032.

[26] D. Hasic, R. Lorbeer, R.C. Bertheau, et al., Vertebral bone marrow fat is independently associated to VAT but not to SAT: KORA FF4—whole-body MR imaging in a population-based cohort, Nutrients 12 (2020) 1527, https://doi.org/10.3390/nu12051527.

[27] Y. Leonhardt, F.T. Gassert, G. Feuerriegel, et al., Vertebral bone marrow T2* mapping using chemical shift encoding-based water-fat separation in the quantitative analysis of lumbar osteoporosis and osteoporotic fractures, Quant. Imaging Med. Surg. 11 (2021) 3715–3725, https://doi.org/10.21037/qims-20-1373.

[28] M. Jung, S. Rospleszcz, M.T. Löffler, et al., Association of lumbar vertebral bone marrow and paraspinal muscle fat composition with intervertebral disc degeneration: 3T quantitative MRI findings from the population-based KORA study, Eur. Radiol. 33 (2023) 1501–1512, https://doi.org/10.1007/s00330-022-09140-4.

[29] T. Kart, M. Fischer, T. Küstner, et al., Deep learning-based automated abdominal organ segmentation in the UK biobank and german national cohort magnetic resonance imaging studies, Invest. Radiol. 56 (2021) 401, https://doi.org/10.1097/RLI.0000000000000755.

[30] T. Haueise, N. Stefan, T.J. Schulz, et al., Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow, Z Für Med. Phys. (2023), https://doi.org/10.1016/j.zemedi.2022.12.004.

[31] H. Yamazaki, S. Tauchi, J. Machann, et al., Fat distribution patterns and future type 2 diabetes, Diabetes 71 (2022) 1937–1945, https://doi.org/10.2337/db22-0315.

[32] J.Z. Bojorquez, S. Bricq, C. Acquitter, et al., What are normal relaxation times of tissues at 3 T? Magn. Reson. Imaging 35 (2017) 69–80, https://doi.org/10.1016/j.mri.2016.08.021.

[33] C.M.J. de Bazelaire, G.D. Duhamel, N.M. Rofsky, D.C. Alsop, MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results, Radiology 230 (2004) 652–659, https://doi.org/10.1148/radiol.2303021331.

[34] F. Schick, H. Bongers, W.I. Jung, et al., Volume-selective proton MRS in vertebral bodies, Magn. Reson. Med. 26 (1992) 207–217, https://doi.org/10.1002/mrm.1910260203.

[35] F. Bamberg, H.-U. Kauczor, S. Weckbach, et al., Whole-body MR imaging in the German national cohort: rationale, design, and technical background, Radiology 277 (2015) 206–220, https://doi.org/10.1148/radiol.2015142272.

[36] F. Bamberg, H. Hetterich, S. Rospleszcz, et al., Subclinical disease burden as assessed by whole-body MRI in subjects with prediabetes, subjects with diabetes, and normal control subjects from the general population: the KORA-MRI study, Diabetes 66 (2017) 158–169, https://doi.org/10.2337/db16-0630.

[37] R.C. Bertheau, R. Lorbeer, J. Nattenmüller, et al., Bone marrow fat fraction assessment in regard to physical activity: KORA FF4–3-T MR imaging in a population-based cohort, Eur. Radiol. 30 (2020) 3417–3428, https://doi.org/10.1007/s00330-019-06612-y.

[38] D.C. Karampinos, S. Ruschke, M. Dieckmeyer, et al., Quantitative MRI and spectroscopy of bone marrow, J. Magn. Reson. Imaging JMRI 47 (2018) 332–353, https://doi.org/10.1002/jmri.25769.

[39] M.A. Bredella, M. Torriani, R.H. Ghomi, et al., Vertebral bone marrow fat is positively associated with visceral fat and inversely associated with IGF-1 in obese women, Obes. Silver Spring Md. 19 (2011) 49–53, https://doi.org/10.1038/oby.2010.106.

[40] F. Schick, H. Einsele, B. Weiß, et al., Assessment of the composition of bone marrow prior to and following autologous BMT and PBSCT by magnetic resonance, Ann. Hematol. 72 (1996) 361–370, https://doi.org/10.1007/s002770050187.

[41] T.J. Littlejohns, J. Holliday, L.M. Gibson, et al., The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions, Nat. Commun. 11 (2020) 2624, https://doi.org/10.1038/s41467-020-15948-9.

[42] D.E. Horng, D. Hernando, S.B. Reeder, Quantification of liver fat in the presence of iron overload, J. Magn. Reson. Imaging JMRI 45 (2017) 428–439, https://doi.org/10.1002/jmri.25382.

[43] H. Wang, Y. Leng, Y. Gong, Bone marrow fat and hematopoiesis, Front. Endocrinol. 9 (2018) 694, https://doi.org/10.3389/fendo.2018.00694.

3 Discussion

Application of advanced MR-based body composition analysis, especially when performed in large cohort studies, has the potential to enhance our understanding of how adipose tissue (AT) distribution affects the onset and progression of metabolic diseases. By grouping of study participants according to their AT volume and distribution to identify body composition phenotypes with distinct allocations of abdominal and ectopic AT depots, this approach can also reveal the prevalence and associated cardiometabolic disease burden of different AT compartments at a population level, aiding in the identification of region-specific, age-specific, or sex-specific characteristics. Furthermore, MR-derived phenotypes of AT distribution can augment current risk prediction models, thereby improving individual risk assessments presumably indicating the necessity for targeted, personalized public health interventions and offering a more detailed and nuanced understanding of risk factors compared to traditional methods (Linge et al. 2024; Agrawal et al. 2023; Niedermayer et al. 2024).

A core methodology is the standardized, reliable, precise, and efficient localization and quantification of single AT compartments using, for example, deep learning (DL)-based image analysis. In summary, the presented results from paper 1 (see. sec. 2.1) demonstrate the effectiveness of applying a DL-based model for automated assessment of volume of visceral AT (VAT) and subcutaneous AT (SAT) in the trunk achieving precision comparable to skilled human observers in a fraction of time. Based on 30 carefully curated and manually annotated examples, i.e., by selection of 15 men and 15 women stratified by age and BMI, we successfully derived a tool for anatomically standardized assessment of VAT and SAT, i.e., from the femoral heads to the cardiac apex or the humeral heads, respectively, to quantify the volumes of VAT and SAT from more than 11,000 participants from the German National Cohort (NAKO), a large cohort of study participants from the general population.

Additionally, evaluation of integrity and plausibility of the results by uncertainty-based outlier detection showed the feasibility of handling data from large cohorts when manual checking of the complete sample is time-consuming, and therefore, not feasible. A further advantage of the pairwise comparative approach was the observation that the segmentations of the model differed strongly mainly in cases where image reconstruction errors (partial fat-water swaps) were present. This allowed further image errors to be identified and excluded from the analysis,

including in the data of the already quality-assured NAKO cohort.

Furthermore, approaching quantification of AT compartments by semantic segmentation allows analyzing the topography, for example, the distribution of VAT and SAT (or their ratio) along the craniocaudal body axis. Our extension of this approach to the quantification of bone marrow adipose tissue (BMAT) can be considered as an exemplary case for the possibility of increasing the level of detail of future body composition analyses on large cohorts. Using standardized shape-independent quantification of proton density fat fraction (PDFF) in vertebral bone marrow, as proposed in paper 2 (see sec. 2.2), additionally provides visualization of subregional, i.e., intravertebral, distribution patterns of vertebral bone marrow PDFF and thereby enables systematic analysis. Future investigations of these distributions will presumably enhance our understanding of vertebral bone marrow functionality, aiding large-scale epidemiological studies and follow-up analyses, for example, in patients with hematological diseases, which affect the composition of red bone marrow, or therapy responses.

Since only two-point Dixon MRI covering the entire spine were acquired in the NAKO, our findings suggest that using two-point Dixon imaging for fat fraction quantification in vertebral bone marrow presents higher systematic errors compared to liver or skeletal muscle. The possibility of their compensation using simple linear models (see sec. 2.3), can be beneficially translated from the DZD studies to the NAKO as acquisition parameters remained constant. This transferability of the results from paper 2 and paper 3 to the NAKO was in fact intended, as data from more than 30,000 participants from the baseline examination were made available by the end of 2023 as well as data from recently completed and ongoing follow-ups are expected in the near future, providing a large sample for detailed analysis of AT distribution and its longitudinal trajectories (Rospleszcz et al. 2019; Niedermayer et al. 2024). Additionally, as standardized imaging protocols in ongoing (multicentric) DZD studies include whole-body Dixon MRI (and are expected to do so for many years), the methods developed in this thesis will further be applied.

However, the results presented are subject to limitations and challenges. First, all segmentation models applied are derived from data in a supervised manner, i.e., there is a need for manually annotated training examples. This task is time-consuming, implicitly subjective in the way delineations between different AT compartments are realized, and therefore a source of noise in model training process, especially when performed by multiple persons to save time, and bias. This is closely related to a second important limitation, namely, the seg-

mentation models are only validated against a very small test data set that was not used during training. Therefore, regular visual inspection of the model outputs is necessary but – once again – biased by subjectivity. Third, all cohorts are subject to a limiting specificity of study populations or conditions: data are acquired from (at least different) scanners from the same manufacturer and not validated for applicability to images from other devices. First tests using data acquired with scanners from different vendors, for example, in multicentric DZD cohorts like PLIS or UKB, provide promising preliminary results and indicate that an application could be possible without further adaptation of the segmentation tools developed in this thesis. A more in-depth evaluation has not yet taken place and is still pending. More disruptive is a look at the included study populations: In the NAKO, a vast majority of participants self-identified as White individuals, in DZD studies race/ethnicity is not even collected but assumed to also be a White majority. With more than 80% of people with future type 2 diabetes living in low- and middle-income countries (GBD 2021 Diabetes Collaborators 2023) and with type 2 diabetes "becoming more and more a disease of inequity" (Jumpertz von Schwartzenberg et al. 2024), more diverse study populations are urgently needed to validate the proposed methodology to be reliably working with imaging data from non-White individuals. Especially in this context, the costs and limited availability of MRI as the underlying imaging method represent a disadvantage of the proposed analysis approaches. Surrogate markers, for example, fat mass ratio derived from DXA, have been linked to cardiometabolic diseases and can help to identify individuals with unhealthy AT distributions without requiring whole-body MRI (Agrawal et al. 2024).

Since the introduction of nnU-Net (Isensee et al. 2021) at the latest, the basic problem of (supervised) medical image segmentation can be considered solved from a methodological and technical perspective, provided that a well-curated training data set is available. Although further optimizations are always possible, for example, by development of time-saving annotation strategies (Gotkowski et al. 2024), acknowledging the above-mentioned fact enables broad applications of the methodology, for example to questions of body composition related to metabolic research, as exemplified in this thesis.

It has been pointed out that there is an innovation bias towards new model architectures in the field (Isensee et al. 2024). However, unreliable evaluations due to inadequately chosen validation metrics (Reinke et al. 2024), many new model architectures fail to prove superiority in comprehensive benchmarking (Isensee et al. 2024). On the other hand, technical developments in the field of DL re-

search, for example, foundations models like GPT-4 (OpenAI et al. 2024) that are highly flexible and able to adapt to many different tasks, and are resource-intensive (i.e., data collection, computational resources, and environmental costs) (Moor et al. 2023) and mainly driven by large actors from industry like OpenAI or Google. In April 2024, the Helmholtz Association announced the Helmholtz Foundation Model Initiative funding, among others, the development of a foundation model trained on nearly five million 3D medical images including MRI, CT and PET (*Helmholtz Invests 23 Million in Research on AI Foundation Models* 2024). Multi-modal foundation models are expected to contribute to clinical use cases, for example, by supporting radiologists throughout their workflow (Zhang et al. 2024).

However, from the application-oriented perspective of this thesis, further developments of the proposed methods will enrich body composition analysis with a more detailed (sub-) segmentation of AT compartments.

3.1 Future work

Foremost, future work includes the extension of the proposed methodology to a more fine-grained analysis of AT compartments, for example, further subdivision of VAT as proposed by Shen et al. (Shen et al. 2003), or by inclusion of additional AT compartments and ectopic fat depots that could have metabolic implications. Due to the proximity of epicardial AT to the coronary arteries and the atrial myocardium, it is thought to influence the development and morphology of atherosclerotic plaques by serving as a mediator of local inflammation (Talman et al. 2014) and likewise linked to several cardiovascular diseases (Iacobellis 2022). In the KORA study, an independent link between epicardial AT and left ventricular dysfunction was demonstrated (Rado et al. 2019). In addition to its association with hypertension, atherosclerosis and chronic kidney disease, elevated renal sinus fat is associated with a distinct subphenotype of prediabetes (Zelicha et al. 2018; Wagner et al. 2021), increased intramuscular fat is associated with a reduction in insulin-stimulated glucose uptake into the muscle cells (insulin resistance of the muscles) (Brøns et al. 2017), and ectopic fat accumulation in the pancreas is related to impaired glucose secretion (Wagner et al. 2022).

In combination with additional metabolic parameters, data from NAKO can be used to further investigate, delineate, and substantiate phenotypes of body composition, for example, the thin-on-the-outside fat-on-the-inside subphenotype (Thomas et al. 2012) or metabolically healthy obesity with high insulin sensitivity

(Stefan 2008). Broadening the perspective to include lifestyle and socioeconomic risk factors could provide further hints to predisposition and pathogenesis and could make a valuable contribution to their prevention or interventional treatment.

Furthermore, AT profiles as proposed by Machann et al. (Machann et al. 2005b) have not yet been considered for further research purposes. A productive extension of body composition phenotyping by inclusion of AT topography would make use of the complete information acquired by 3D whole-body MRI. For example, paper 1 revealed sex-, age-, and BMI-related differences in craniocaudal distribution of VAT and SAT.

As outlined by Gupta, detailed body composition analysis can also help to further investigate the relationship of (visceral) adiposity and the pathogenesis and prognosis of various cancers which "is evident, yet still poorly understood" (Gupta et al. 2024). Results from NAKO might help to uncover this relationship.

Sneed and Morrison concluded in a recent review that "implementation of body composition methods into healthcare settings can augment clinical decision making and future studies should focus on the use of these techniques in areas of clinical research." (Sneed et al. 2021) From a clinical and individual-based perspective, the applicability of advanced Dixon MRI in daily routine and MR-based body composition analysis are subject to time- and cost constraints. Using sonography for the measurement of abdominal AT thickness could be a more feasible alternative saving time and offering the possibility of carrying out a risk classification with simpler means. Validity and reproducibility have already been documented in the pretest of the NAKO (Schlecht et al. 2014; Schlecht et al. 2016) and gold standard results obtained from the methodology presented can be used in further analyses on a large sample aiming at deriving possible reference values (B Fischer et al. 2020). Derivation of simpler and more accessible markers or the development of methods for faster assessment of AT distribution could help tackling the "disease of inequity" (Jumpertz von Schwartzenberg et al. 2024).

However, MRI can also provide additional insights in the relationship between brown AT that is, independent of BMI, inversely correlated with risk for developing various cardiovascular disorders (Gupta et al. 2024; Becher et al. 2021). MR spectroscopy has recently been used to distinguish brown AT from white AT, as in brown AT the main methylene resonance had shorter T2 relaxation times and lower proton densities of unsaturated and polyunsaturated fat, presumably originating from lower fatty acid fluidity or increased iron content found in the mitochondrial cells present in brown AT (Ouwerkerk et al. 2021). Imaging-based analysis of fatty acid composition (Schneider et al. 2019) is an emerging research

topic and a potentially useful technique in providing additional information from MR-based body composition analysis. Another “missing variable in many of the earlier discussions of adiposity was the health, or quality, of adipose tissue”, potentially being “as important as adipose tissue mass and location as a predictor and determinant of metabolic health” (Gupta et al. 2024) and could be addressed by imaging inflammation and fibrosis (Vishvanath et al. 2019; A Fischer et al. 2021).

Finally, aiming for collection of data of a diverse population and its intersectional examination should be a major part of future research (Harari et al. 2021). For example, paper 3 found significant sex differences in T2* within vertebral BMAT and therefore suggests the necessity for sex-specific image reconstruction and postprocessing.

3.2 Conclusions

Standardized MR-based localization and quantification of adipose tissue compartments are useful for the analysis of large-scale epidemiological studies allowing for accurate interindividual comparisons and spatial analysis of tissues, particularly in the context of varying anatomical and physiological conditions.

4 Summary

Obesity is a global public health challenge and a heterogeneous chronic disease associated with disorders of, for example, the endocrine and cardiovascular system. In 2022, 2.5 billion adults were overweight, of which 890 million were living with obesity; the global prevalence has almost tripled since 1975. Starting in the 1950s, research has shown that the distribution of adipose tissue in the body plays a crucial role in characterizing different phenotypes of obesity and that subphenotypes of abdominal obesity are associated with metabolic diseases. This necessitates a closer examination of the distribution of adipose tissue. Non-invasive imaging techniques such as magnetic resonance imaging (MRI) are particularly suitable for this purpose. They are also used in large cohort studies such as NAKO, aiming to investigate the causes of chronic population-wide diseases. Retrospective analyses of large cohort studies as well as the consecutive evaluation of adipose tissue compartments from MRI in ongoing cross-sectional and intervention studies are dependent on standardized, reliable, precise, objective and time-saving methods for image analysis.

The aim of this work is to contribute to the standardization of MR-based localization and quantification of adipose tissue compartments by adapting and evaluating an existing deep learning-based image segmentation tool (nnU-Net) for the standardized quantification of visceral and subcutaneous adipose tissue and to demonstrate its applicability on image datasets of more than 11,000 participants of the NAKO. Furthermore, a method for the standardized evaluation of the fat fraction and its subregional distribution in the hematopoietic bone marrow of the vertebral bodies is developed. In addition, a method for correcting confounded measurements of fat fraction in the liver, muscles and bone marrow of the vertebral bodies is proposed.

The results show that the developed segmentation model can be used effectively for automatic localization and quantification as well as for characterizing the distribution of visceral and subcutaneous adipose tissue along the body axis. Furthermore, it was shown that the developed method for analyzing the fat fraction in bone marrow provides robust and precise results and enables the investigation of group-specific distribution patterns within the vertebral bodies. In addition, the comparison of commercially available two-point and multi-echo Dixon MRI in bone marrow shows higher systematic errors in the quantification of fat fraction compared to liver and muscle. This error can be corrected using linear models

derived from the data.

The presented approaches for standardized MR-based analysis of body composition and their applications in a large population-based cohort study contribute to a better understanding of the influence of adipose tissue distribution on metabolic diseases, allow the identification of specific phenotypes associated with cardiometabolic risk and thus provide an essential contribution to risk stratification. In future work, the proposed methods for adipose tissue quantification are to be applied to the more than 30,000 MR image data sets of the NAKO. Furthermore, the methodology is to be extended to additional (ectopic) fat compartments and used, for example, to quantify epi-/pericardial adipose tissue or the fat fraction in the pancreas.

5 German summary

Adipositas ist eine globale Herausforderung für die allgemeine Gesundheit der Bevölkerung und eine heterogene chronische Krankheit, die mit Störungen z. B. des endokrinen oder kardiovaskulären Systems einhergeht. Im Jahr 2022 waren 2,5 Milliarden Erwachsene übergewichtig, von denen 890 Millionen mit Adipositas lebten; die weltweite Prävalenz hat sich seit 1975 fast verdreifacht. Seit den 1950er-Jahren zeigen Forschungsergebnisse, dass die Verteilung des Fettgewebes im Körper eine entscheidende Rolle für die Charakterisierung von verschiedenen Phänotypen der Adipositas spielt und Subphänotypen von abdomineller Adipositas mit Stoffwechselerkrankungen in Verbindung stehen. Dies führt zu der Notwendigkeit, die Verteilung von Fettgewebe genauer zu untersuchen. Dazu eignen sich besonders nicht-invasive bildgebende Verfahren wie die Magnetresonanztomographie (MRT), die auch in großen Kohortenstudien, z. B. der NAKO Gesundheitsstudie zur Erforschung der Ursachen chronischer bevölkerungsweiter Krankheiten, zum Einsatz kommt. Retrospektive Analysen großer Kohortenstudien sowie die konsekutive Auswertung von Fettgewebskompartimenten aus der MRT in laufenden Querschnitts- und Interventionsstudien sind auf standardisierte, zuverlässige, präzise, objektive und zeitsparende Methoden zur Bildanalyse angewiesen.

Das Ziel dieser Arbeit ist es, einen Beitrag zur Standardisierung der MR-basierten Lokalisierung und Quantifizierung von Fettgewebskompartimenten zu leisten, indem ein auf Deep Learning basierendes Bildsegmentierungswerkzeug (nnU-Net) für die standardisierte Quantifizierung von viszeralem und subkutanem Fettgewebe entwickelt und evaluiert wird, und dessen Anwendbarkeit auf Bilddatensätzen von mehr als 11.000 Teilnehmenden der NAKO gezeigt wird. Weiterhin wird eine Methode zur standardisierten Auswertung der Fettanteils und dessen subregionaler Verteilung im blutbildenden Knochenmark der Wirbelkörper entwickelt. Außerdem wird eine Möglichkeit zur Korrektur von störgrößenbehafteten Messungen des Fettanteils in Leber, Muskulatur und Knochenmark der Wirbelkörper vorgeschlagen.

Die Ergebnisse zeigen, dass das entwickelte Segmentierungsmodell sowohl zur automatischen Lokalisierung und Quantifizierung als auch zur Charakterisierung der Verteilung von viszeralem und subkutanem Fett entlang der Körperachse effektiv eingesetzt werden kann. Weiterhin wurde gezeigt, dass die entwickelte Methode zur Analyse des Fettanteils im Knochenmark robuste und präzise Er-

gebnisse liefert und die Untersuchung gruppenspezifischer Verteilungsmuster innerhalb der Wirbelkörper ermöglicht. Außerdem zeigt der Vergleich von kommerziell verfügbaren zweipunkt- und multiecho-Dixon-Methoden im Knochenmark höhere systematische Fehler in der Quantifizierung des Fettanteils im Vergleich zur Leber bzw. Muskulatur. Mittels aus den Daten abgeleiteter linearer Modelle lässt sich dieser Fehler korrigieren.

Die vorgestellten Ansätze zur standardisierten MR-basierten Analyse der Körperzusammensetzung und deren Anwendungen in einer großen, bevölkerungsbasierten Kohortenstudie tragen zum besseren Verständnis des Einflusses der Fettverteilung auf metabolische Erkrankungen bei, ermöglichen die Identifizierung spezifischer Phänotypen, die mit einem kardiometabolischen Risiko verbunden sind und liefern somit einen essentiellen Beitrag zur Risikostratifizierung. In zukünftigen Arbeiten sollen die vorgeschlagenen Methoden zur Fettquantifizierung auf die insgesamt mehr als 30.000 MR-Bilddatensätze der NAKO angewendet werden. Außerdem soll die Methodik auf zusätzliche (ektope) Fettkompartimente ausgeweitet werden, und beispielsweise zur Quantifizierung von epi-/perikardialem Fett oder dem Fettanteil in der Bauchspeicheldrüse eingesetzt werden.

Bibliography

Aaron N, Costa S, Rosen CJ, Qiang L. The Implications of Bone Marrow Adipose Tissue on Inflammaging. *Front Endocrinol (Lausanne)*. 2022;13:853765. DOI: 10.3389/fendo.2022.853765.

Agrawal S, Klarqvist MDR, Diamant N, et al. *Association of Machine Learning-Derived Measures of Body Fat Distribution in >40,000 Individuals with Cardiometabolic Diseases*. 2021. DOI: 10.1101/2021.05.07.21256854.

Agrawal S, Klarqvist MDR, Diamant N, et al. BMI-adjusted Adipose Tissue Volumes Exhibit Depot-Specific and Divergent Associations with Cardiometabolic Diseases. *Nat Commun*. 2023;14(1):266. DOI: 10.1038/s41467-022-35704-5.

Agrawal S, Luan J, Cummings BB, Weiss EJ, Wareham NJ, Khera AV. Relationship of Fat Mass Ratio, a Biomarker for Lipodystrophy, With Cardiometabolic Traits. *Diabetes*. 2024;73(7):1099–1111. DOI: 10.2337/db23-0575.

Ahrens W, Greiser H, Linseisen J, et al. [The design of a nationwide cohort study in Germany : the pretest studies of the German National Cohort (GNC)]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2014;57(11):1246–1254. DOI: 10.1007/s00103-014-2042-0.

Ambrosi TH, Scialdone A, Graja A, et al. Adipocyte Accumulation in the Bone Marrow during Obesity and Aging Impairs Stem Cell-Based Hematopoietic and Bone Regeneration. *Cell Stem Cell*. 2017;20(6):771–784.e6. DOI: 10.1016/j.stem.2017.02.009.

Aoki T, Yamaguchi S, Kinoshita S, Hayashida Y, Korogi Y. Quantification of Bone Marrow Fat Content Using Iterative Decomposition of Water and Fat with Echo Asymmetry and Least-Squares Estimation (IDEAL): Reproducibility, Site Variation and Correlation with Age and Menopause. *Br J Radiol*. 2016;89(1065):20150538. DOI: 10.1259/bjr.20150538.

Bamberg F, Hetterich H, Rospleszcz S, et al. Subclinical Disease Burden as Assessed by Whole-Body MRI in Subjects With Prediabetes, Subjects With Diabetes, and Normal Control Subjects From the General Population: The KORA-MRI Study. *Diabetes*. 2016;66(1):158–169. DOI: 10.2337/db16-0630.

Bamberg F, Kauczor HU, Weckbach S, et al. Whole-Body MR Imaging in the German National Cohort: Rationale, Design, and Technical Background. *Radiology*. 2015;277(1):206–220. DOI: 10.1148/radiol.2015142272.

Bashir MR, Zhong X, Nickel MD, et al. Quantification of Hepatic Steatosis with a Multistep Adaptive Fitting MRI Approach: Prospective Validation against MR Spectroscopy. *AJR Am J Roentgenol*. 2015;204(2):297–306. DOI: 10.2214/AJR.14.12457.

Baum T, Rohrmeier A, Syväri J, et al. Anatomical Variation of Age-Related Changes in Vertebral Bone Marrow Composition Using Chemical Shift Encoding-Based Water-Fat Magnetic Resonance Imaging. *Front Endocrinol (Lausanne)*. 2018;9:141. DOI: 10.3389/fendo.2018.00141.

Baum T, Yap SP, Dieckmeyer M, et al. Assessment of Whole Spine Vertebral Bone Marrow Fat Using Chemical Shift-Encoding Based Water-Fat MRI. *J Magn Reson Imaging*. 2015;42(4):1018–1023. DOI: 10.1002/jmri.24854.

Becher T, Palanisamy S, Kramer DJ, et al. Brown Adipose Tissue Is Associated with Cardiometabolic Health. *Nat Med*. 2021;27(1):58–65. DOI: 10.1038/s41591-020-1126-7. (Visited on 07/31/2024).

Beekman KM, Duque G, Corsi A, Tencerova M, Bisschop PH, Paccou J. Osteoporosis and Bone Marrow Adipose Tissue. *Curr Osteoporos Rep*. 2023;21(1):45–55. DOI: 10.1007/s11914-022-00768-1.

Behrens G, Gredner T, Stock C, Leitzmann MF, Brenner H, Mons U. Cancers Due to Excess Weight, Low Physical Activity, and Unhealthy Diet. *Dtsch Arztebl Int*. 2018;115(35-36):578–585. DOI: 10.3238/arztebl.2018.0578.

Bertheau RC, Lorbeer R, Nattenmüller J, et al. Bone Marrow Fat Fraction Assessment in Regard to Physical Activity: KORA FF4–3-T MR Imaging in a Population-Based Cohort. *Eur Radiol*. 2020;30(6):3417–3428. DOI: 10.1007/s00330-019-06612-y.

Bhanu PK, Arvind CS, Yeow LY, Chen WX, Lim WS, Tan CH. CAFT: A Deep Learning-Based Comprehensive Abdominal Fat Analysis Tool for Large Cohort Studies. *Magn Reson Mater Phy*. 2022;35(2):205–220. DOI: 10.1007/s10334-021-00946-9.

Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What Are Normal Relaxation Times of Tissues at 3 T? *Magn Reson Imaging*. 2017;35:69–80. DOI: 10.1016/j.mri.2016.08.021.

Bolan PJ, Arentsen L, Sueblinvong T, et al. Water-Fat MRI for Assessing Changes in Bone Marrow Composition Due to Radiation and Chemotherapy in Gynecologic Cancer Patients. *J Magn Reson Imaging*. 2013;38(6):1578–1584. DOI: 10.1002/jmri.24071.

Borga M. MRI Adipose Tissue and Muscle Composition Analysis-a Review of Automation Techniques. *Br J Radiol*. 2018;91(1089):20180252. DOI: 10.1259/bjr.20180252.

Borga M, Thomas EL, Romu T, et al. Validation of a Fast Method for Quantification of Intra-Abdominal and Subcutaneous Adipose Tissue for Large-Scale Human Studies. *NMR Biomed.* 2015;28(12):1747–1753. DOI: 10.1002/nbm.3432. (Visited on 06/03/2024).

Borga M, West J, Bell JD, et al. Advanced Body Composition Assessment: From Body Mass Index to Body Composition Profiling. *J Investig Med*. 2018;66(5):1–9. DOI: 10.1136/jim-2018-000722.

Borkan GA, Gerzof SG, Robbins AH, Hults DE, Silbert CK, Silbert JE. Assessment of Abdominal Fat Content by Computed Tomography. *Am J Clin Nutr*. 1982;36(1):172–177. DOI: 10.1093/ajcn/36.1.172.

Bray GA. Beyond BMI. *Nutrients*. 2023;15(10):2254. DOI: 10.3390/nu15102254.

Bredella MA, Gill CM, Gerweck AV, et al. Ectopic and Serum Lipid Levels Are Positively Associated with Bone Marrow Fat in Obesity. *Radiology*. 2013;269(2):534–541. DOI: 10.1148/radiology.13130375.

Bredella MA, Torriani M, Ghomi RH, et al. Vertebral Bone Marrow Fat Is Positively Associated with Visceral Fat and Inversely Associated with IGF-1 in Obese Women. *Obesity (Silver Spring)*. 2011;19(1):49–53. DOI: 10.1038/oby.2010.106.

Britton KA, Fox CS. Ectopic Fat Depots and Cardiovascular Disease. *Circulation*. 2011;124(24). DOI: 10.1161/CIRCULATIONAHA.111.077602.

Brøns C, Grunnet LG. Mechanisms In Endocrinology: Skeletal Muscle Lipotoxicity in Insulin Resistance and Type 2 Diabetes: A Causal Mechanism or an Innocent Bystander? *Eur J Endocrinol*. 2017;176(2):R67–R78. DOI: 10.1530/EJE-16-0488.

Burakiewicz J, Sinclair CDJ, Fischer D, Walter GA, Kan HE, Hollingsworth KG. Quantifying Fat Replacement of Muscle by Quantitative MRI in Muscular Dystrophy. *J Neurol*. 2017;264(10):2053–2067. DOI: 10.1007/s00415-017-8547-3.

Burian E, Rohrmeier A, Schlaeger S, et al. Lumbar Muscle and Vertebral Bodies Segmentation of Chemical Shift Encoding-Based Water-Fat MRI: The Reference Database MyoSegmenTUM Spine. *BMC Musculoskelet Disord*. 2019;20(1):152. DOI: 10.1186/s12891-019-2528-x.

Bydder M, Yokoo T, Hamilton G, et al. Relaxation Effects in the Quantification of Fat Using Gradient Echo Imaging. *Magnetic Resonance Imaging*. 2008;26(3):347–359. DOI: 10.1016/j.mri.2007.08.012.

Carmona R, Pritz J, Bydder M, et al. Fat Composition Changes in Bone Marrow during Chemotherapy and Radiation Therapy. *Int J Radiat Oncol Biol Phys*. 2014;90(1):155–163. DOI: 10.1016/j.ijrobp.2014.05.041.

Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *Med. Image Comput. Comput.-Assist. Interv. – MICCAI 2016*. Ed. by S Ourselin, L Joskowicz, MR Sabuncu, G Unal, W Wells. Cham: Springer International Publishing, 2016: 424–432. DOI: 10.1007/978-3-319-46723-8_49.

Cypess AM. Reassessing Human Adipose Tissue. *N Engl J Med*. 2022;386(8):768–779. DOI: 10.1056/NEJMra2032804. (Visited on 05/31/2024).

de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR Imaging Relaxation Times of Abdominal and Pelvic Tissues Measured in Vivo at 3.0 T: Preliminary Results. *Radiology*. 2004;230(3):652–659. DOI: 10.1148/radiol.2303021331.

Després JP. Body Fat Distribution and Risk of Cardiovascular Disease: An Update. *Circulation*. 2012;126(10):1301–1313. DOI: 10.1161/CIRCULATIONAHA.111.067264.

Després JP, Lemieux I. Abdominal Obesity and Metabolic Syndrome. *Nature*. 2006;444(7121):881–887. DOI: 10.1038/nature05488.

Després JP, Nadeau A, Tremblay A, et al. Role of Deep Abdominal Fat in the Association Between Regional Adipose Tissue Distribution and Glucose Tolerance in Obese Women. *Diabetes*. 1989;38(3):304–309. DOI: 10.2337/diab.38.3.304. (Visited on 07/17/2024).

Dieckmeyer M, Junker D, Ruschke S, et al. Vertebral Bone Marrow Heterogeneity Using Texture Analysis of Chemical Shift Encoding-Based MRI: Variations in Age, Sex, and Anatomical Location. *Front Endocrinol (Lausanne)*. 2020;11:555931. DOI: 10.3389/fendo.2020.555931.

Dixon WT. Simple Proton Spectroscopic Imaging. *Radiology*. 1984;153(1):189–194. DOI: 10.1148/radiology.153.1.6089263.

Donohoe CL, Doyle SL, Reynolds JV. Visceral Adiposity, Insulin Resistance and Cancer Risk. *Diabetol Metab Syndr*. 2011;3:12. DOI: 10.1186/1758-5996-3-12.

Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C. The Importance of Skip Connections in Biomedical Image Segmentation. *Deep Learn. Data*

Labeling Med. Appl. Ed. by G Carneiro, D Mateus, L Peter, et al. Cham: Springer International Publishing, 2016: 179–187. DOI: 10.1007/978-3-319-46976-8_19.

Eggers H, Börnert P. Chemical Shift Encoding-Based Water-Fat Separation Methods. *J Magn Reson Imaging*. 2014;40(2):251–268. DOI: 10.1002/jmri.24568.

Engelke K, Chaudry O, Gast L, et al. Magnetic Resonance Imaging Techniques for the Quantitative Analysis of Skeletal Muscle: State of the Art. *J Orthop Translat*. 2023;42:57–72. DOI: 10.1016/j.jot.2023.07.005.

Estrada S, Lu R, Conjeti S, et al. FatSegNet: A Fully Automated Deep Learning Pipeline for Adipose Tissue Segmentation on Abdominal Dixon MRI. *Magn Reson Med*. 2020;83(4):1471–1483. DOI: 10.1002/mrm.28022.

Evans DJ, Hoffmann RG, Kalkhoff RK, Kissebah AH. Relationship of Body Fat Topography to Insulin Sensitivity and Metabolic Profiles in Premenopausal Women. *Metabolism*. 1984;33(1):68–75. DOI: 10.1016/0026-0495(84)90164-1.

Falank C, Fairfield H, Reagan MR. Signaling Interplay between Bone Marrow Adipose Tissue and Multiple Myeloma Cells. *Front Endocrinol (Lausanne)*. 2016;7:67. DOI: 10.3389/fendo.2016.00067.

Fischer A, Schick F. Towards Detection of Inflammation in Adipose Tissue: Microscopic Field Simulations to Estimate Water Signal Properties. *Z Med Phys*. 2021;31(4):394–402. DOI: 10.1016/j.zemedi.2021.03.002.

Fischer B, Sedlmeier AM, Hartwig S, et al. [Anthropometric measures in the German National Cohort-more than weight and height]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2020;63(3):290–300. DOI: 10.1007/s00103-020-03096-w.

Fritsche A, Wagner R, Heni M, et al. Different Effects of Lifestyle Intervention in High- and Low-Risk Prediabetes: Results of the Randomized Controlled Prediabetes Lifestyle Intervention Study (PLIS). *Diabetes*. 2021;70(12):2785–2795. DOI: 10.2337/db21-0526.

Gallagher D, Kelley DE, Yim JE, et al. Adipose Tissue Distribution Is Different in Type 2 Diabetes. *Am J Clin Nutr*. 2009;89(3):807–814. DOI: 10.3945/ajcn.2008.26955.

Gassert FT, Kufner A, Gassert FG, et al. MR-based Proton Density Fat Fraction (PDFF) of the Vertebral Bone Marrow Differentiates between Patients with and without Osteoporotic Vertebral Fractures. *Osteoporos Int*. 2022;33(2):487–496. DOI: 10.1007/s00198-021-06147-3.

GBD 2021 Diabetes Collaborators. Global, Regional, and National Burden of Diabetes from 1990 to 2021, with Projections of Prevalence to 2050: A Systematic Analysis for the Global Burden of Disease Study 2021. *Lancet*. 2023;402(10397):203–234. DOI: 10.1016/S0140-6736(23)01301-6.

German National Cohort (GNC) Consortium. The German National Cohort: Aims, Study Design and Organization. *Eur J Epidemiol*. 2014;29(5):371–382. DOI: 10.1007/s10654-014-9890-7.

Goodpaster BH, Bergman BC, Brennan AM, Sparks LM. Intermuscular Adipose Tissue in Metabolic Disease. *Nat Rev Endocrinol*. 2023;19(5):285–298. DOI: 10.1038/s41574-022-00784-2.

Gotkowski K, Lüth C, Jäger PF, et al. *Embarrassingly Simple Scribble Supervision for 3D Medical Segmentation*. 2024. DOI: 10.48550/arXiv.2403.12834.

Griffith JF, Yeung DKW, Antonio GE, et al. Vertebral Bone Mineral Density, Marrow Perfusion, and Fat Content in Healthy Men and Men with Osteoporosis: Dynamic Contrast-Enhanced MR Imaging and MR Spectroscopy. *Radiology*. 2005;236(3):945–951. DOI: 10.1148/radiol.2363041425.

Griffith JF, Yeung DKW, Ma HT, Leung JCS, Kwok TCY, Leung PC. Bone Marrow Fat Content in the Elderly: A Reversal of Sex Difference Seen in Younger Subjects. *J Magn Reson Imaging*. 2012;36(1):225–230. DOI: 10.1002/jmri.23619.

Gupta OT, Gupta RK. The Expanding Problem of Regional Adiposity: Revisiting a 1985 Diabetes Classic by Ohlson et Al. *Diabetes*. 2024;73(5):649–652. DOI: 10.2337/dbi24-0021.

Harari L, Lee C. Intersectionality in Quantitative Health Disparities Research: A Systematic Review of Challenges and Limitations in Empirical Studies. *Soc Sci Med*. 2021;277:113876. DOI: 10.1016/j.socscimed.2021.113876.

Hasic D, Lorbeer R, Bertheau RC, et al. Vertebral Bone Marrow Fat Is Independently Associated to VAT but Not to SAT: KORA FF4—Whole-Body MR Imaging in a Population-Based Cohort. *Nutrients*. 2020;12(5):1527. DOI: 10.3390/nu12051527.

Haueise T, Schick F, Stefan N, et al. Analysis of Volume and Topography of Adipose Tissue in the Trunk: Results of MRI of 11,141 Participants in the German National Cohort. *Sci Adv*. 2023;9(19):eadd0433. DOI: 10.1126/sciadv.add0433.

Hayashi T, Fukuzawa K, Yamazaki H, et al. Multicenter, Multivendor Phantom Study to Validate Proton Density Fat Fraction and T2* Values Calculated Using Vendor-Provided 6-Point DIXON Methods. *Clin Imaging*. 2018;51:38–42. DOI: 10.1016/j.clinimag.2018.01.011.

Hegenscheid K, Kühn JP, Völzke H, Biffar R, Hosten N, Puls R. Whole-Body Magnetic Resonance Imaging of Healthy Volunteers: Pilot Study Results from the Population-Based SHIP Study. *Rofo*. 2009;181(8):748–759. DOI: 10.1055/s-0028-1109510.

Heimann T, Meinzer HP. Statistical Shape Models for 3D Medical Image Segmentation: A Review. *Med Image Anal*. 2009;13(4):543–563. DOI: 10.1016/j.media.2009.05.004.

Helmholtz Invests 23 Million in Research on AI Foundation Models. https://www.helmholtz.de/en/newsroom/article/helmholtz-invests-23-million-in-research-on-ai-foundation-models/. 2024. (Visited on 06/17/2024).

Heni M, Machann J, Staiger H, et al. Pancreatic Fat Is Negatively Associated with Insulin Secretion in Individuals with Impaired Fasting Glucose and/or Impaired Glucose Tolerance: A Nuclear Magnetic Resonance Study. *Diabetes Metab Res Rev*. 2010;26(3):200–205. DOI: 10.1002/dmrr.1073.

Hernando D, Liang ZP, Kellman P. Chemical Shift-Based Water/Fat Separation: A Comparison of Signal Models. *Magn Reson Med*. 2010;64(3):811–822. DOI: 10.1002/mrm.22455.

Horng DE, Hernando D, Reeder SB. Quantification of Liver Fat in the Presence of Iron Overload. *J Magn Reson Imaging*. 2017;45(2):428–439. DOI: 10.1002/jmri.25382.

Hu HH, Yokoo T, Bashir MR, et al. Linearity and Bias of Proton Density Fat Fraction as a Quantitative Imaging Biomarker: A Multicenter, Multiplatform, Multivendor Phantom Study. *Radiology*. 2021;298(3):640–651. DOI: 10.1148/radiol.2021202912.

Iacobellis G. Epicardial Adipose Tissue in Contemporary Cardiology. *Nat Rev Cardiol*. 2022;19(9):593–606. DOI: 10.1038/s41569-022-00679-9.

Iglesias JE, Sabuncu MR. Multi-Atlas Segmentation of Biomedical Images: A Survey. *Med Image Anal*. 2015;24(1):205–219. DOI: 10.1016/j.media.2015.06.012.

Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation. *Nat Methods*. 2021;18(2):203–211. DOI: 10.1038/s41592-020-01008-z.

Isensee F, Wald T, Ulrich C, et al. *nnU-Net Revisited: A Call for Rigorous Validation in 3D Medical Image Segmentation*. 2024. DOI: 10.48550/arXiv.2404.09556.

Ismail AA, Cooper C, Felsenberg D, et al. Number and Type of Vertebral Deformities: Epidemiological Characteristics and Relation to Back Pain and Height Loss. *Osteoporos Int*. 1999;9(3):206–213. DOI: 10.1007/s001980050138.

Jang JK, Lee SS, Kim B, et al. Agreement and Reproducibility of Proton Density Fat Fraction Measurements Using Commercial MR Sequences Across Different Platforms: A Multivendor, Multi-Institutional Phantom Experiment. *Invest. Radiol.* 2019;54(8):517. DOI: 10.1097/RLI.0000000000000561.

Jensen MD, Kanaley JA, Reed JE, Sheedy PF. Measurement of Abdominal and Visceral Fat with Computed Tomography and Dual-Energy x-Ray Absorptiometry. *Am J Clin Nutr*. 1995;61(2):274–278. DOI: 10.1093/ajcn/61.2.274.

Jumpertz von Schwartzenberg R, Vazquez Arreola E, Sandforth A, Hanson RL, Birkenfeld AL. Role of Weight Loss-Induced Prediabetes Remission in the Prevention of Type 2 Diabetes: Time to Improve Diabetes Prevention. *Diabetologia*. 2024;67(8):1714–1718. DOI: 10.1007/s00125-024-06178-5. (Visited on 01/29/2025).

Jung M, Rospleszcz S, Löffler MT, et al. Association of Lumbar Vertebral Bone Marrow and Paraspinal Muscle Fat Composition with Intervertebral Disc Degeneration: 3T Quantitative MRI Findings from the Population-Based KORA Study. *Eur Radiol*. 2023;33(3):1501–1512. DOI: 10.1007/s00330-022-09140-4.

Karampinos DC, Melkus G, Baum T, Bauer JS, Rummeny EJ, Krug R. Bone Marrow Fat Quantification in the Presence of Trabecular Bone: Initial Comparison between Water-Fat Imaging and Single-Voxel MRS. *Magn Reson Med*. 2014;71(3):1158–1165. DOI: 10.1002/mrm.24775.

Karampinos DC, Ruschke S, Dieckmeyer M, et al. Modeling of T2* Decay in Vertebral Bone Marrow Fat Quantification. *NMR Biomed*. 2015;28(11):1535–1542. DOI: 10.1002/nbm.3420.

Karampinos DC, Ruschke S, Dieckmeyer M, et al. Quantitative MRI and Spectroscopy of Bone Marrow. *J Magn Reson Imaging*. 2018;47(2):332–353. DOI: 10.1002/jmri.25769.

Karlsson A, Rosander J, Romu T, et al. Automatic and Quantitative Assessment of Regional Muscle Volume by Multi-Atlas Segmentation Using Whole-Body

Water-Fat MRI. *J Magn Reson Imaging*. 2015;41(6):1558–1569. DOI: 10.1002/jmri.24726.

Kart T, Fischer M, Küstner T, et al. Deep Learning-Based Automated Abdominal Organ Segmentation in the UK Biobank and German National Cohort Magnetic Resonance Imaging Studies. *Invest. Radiol.* 2021;56(6):401. DOI: 10.1097/RLI.0000000000000755.

Kiefer LS, Fabian J, Rospleszcz S, et al. Assessment of the Degree of Abdominal Myosteatosis by Magnetic Resonance Imaging in Subjects with Diabetes, Prediabetes and Healthy Controls from the General Population. *European Journal of Radiology*. 2018;105:261–268. DOI: 10.1016/j.ejrad.2018.06.023.

Kiefer LS, Fabian J, Rospleszcz S, et al. Distribution Patterns of Intramyocellular and Extramyocellular Fat by Magnetic Resonance Imaging in Subjects with Diabetes, Prediabetes and Normoglycaemic Controls. *Diabetes Obes Metab*. 2021;23(8):1868–1878. DOI: 10.1111/dom.14413.

Kissebah AH, Vydelingum N, Murray R, et al. Relation of Body Fat Distribution to Metabolic Complications of Obesity. *J Clin Endocrinol Metab*. 1982;54(2):254–260. DOI: 10.1210/jcem-54-2-254.

Krotkiewski M, Björntorp P, Sjöström L, Smith U. Impact of Obesity on Metabolism in Men and Women. Importance of Regional Adipose Tissue Distribution. *J Clin Invest*. 1983;72(3):1150–1162. DOI: 10.1172/JCI111040.

Küstner T, Hepp T, Fischer M, et al. Fully Automated and Standardized Segmentation of Adipose Tissue Compartments via Deep Learning in 3D Whole-Body MRI of Epidemiologic Cohort Studies. *Radiol Artif Intell*. 2020;2(6):e200010. DOI: 10.1148/ryai.2020200010.

Kwok WE. Basic Principles of and Practical Guide to Clinical MRI Radiofrequency Coils. *RadioGraphics*. 2022;42(3):898–918. DOI: 10.1148/rg.210110.

Langner T, Hedström A, Mörwald K, et al. Fully Convolutional Networks for Automated Segmentation of Abdominal Adipose Tissue Depots in Multicenter Water-Fat MRI. *Magn Reson Med*. 2019;81(4):2736–2745. DOI: 10.1002/mrm.27550.

Lapidus L, Bengtsson C, Larsson B, Pennert K, Rybo E, Sjöström L. Distribution of Adipose Tissue and Risk of Cardiovascular Disease and Death: A 12 Year Follow up of Participants in the Population Study of Women in Gothenburg, Sweden. *Br Med J (Clin Res Ed)*. 1984;289(6454):1257–1261. DOI: 10.1136/bmj.289.6454.1257.

Lavynenko O, Abdul-Ghani M, Alatrach M, et al. Combination Therapy with Pioglitazone/Exenatide/Metformin Reduces the Prevalence of Hepatic Fibrosis and Steatosis: The Efficacy and Durability of Initial Combination Therapy for Type 2 Diabetes (EDICT). *Diabetes Obes Metab*. 2022;24(5):899–907. DOI: 10.1111/dom.14650.

Le Ster C, Gambarota G, Lasbleiz J, Guillin R, Decaux O, Saint-Jalmes H. Breath-Hold MR Measurements of Fat Fraction, T1 , and T2 * of Water and Fat in Vertebral Bone Marrow. *J Magn Reson Imaging*. 2016;44(3):549–555. DOI: 10.1002/jmri.25205.

Lean ME, Leslie WS, Barnes AC, et al. Primary Care-Led Weight Management for Remission of Type 2 Diabetes (DiRECT): An Open-Label, Cluster-Randomised Trial. *Lancet*. 2018;391(10120):541–551. DOI: 10.1016/S0140-6736(17)33102-1.

Lean ME, Leslie WS, Barnes AC, et al. 5-Year Follow-up of the Randomised Diabetes Remission Clinical Trial (DiRECT) of Continued Support for Weight Loss Maintenance in the UK: An Extension Study. *Lancet Diabetes Endocrinol*. 2024;12(4):233–246. DOI: 10.1016/S2213-8587(23)00385-6.

Lee DH, Keum N, Hu FB, et al. Development and Validation of Anthropometric Prediction Equations for Lean Body Mass, Fat Mass and Percent Fat in Adults Using the National Health and Nutrition Examination Survey (NHANES) 1999-2006. *Br J Nutr*. 2017;118(10):858–866. DOI: 10.1017/S0007114517002665.

Leonhardt Y, Gassert FT, Feuerriegel G, et al. Vertebral Bone Marrow T2* Mapping Using Chemical Shift Encoding-Based Water-Fat Separation in the Quantitative Analysis of Lumbar Osteoporosis and Osteoporotic Fractures. *Quant Imaging Med Surg*. 2021;11(8):3715–3725. DOI: 10.21037/qims-20-1373.

Li G, Xu Z, Gu H, et al. Comparison of Chemical Shift-Encoded Water-Fat MRI and MR Spectroscopy in Quantification of Marrow Fat in Postmenopausal Females. *J Magn Reson Imaging*. 2017;45(1):66–73. DOI: 10.1002/jmri.25351.

Li X, Kuo D, Schafer AL, et al. Quantification of Vertebral Bone Marrow Fat Content Using 3 Tesla MR Spectroscopy: Reproducibility, Vertebral Variation, and Applications in Osteoporosis. *J Magn Reson Imaging*. 2011;33(4):974–979. DOI: 10.1002/jmri.22489.

Linge J, Borga M, West J, et al. Body Composition Profiling in the UK Biobank Imaging Study. *Obesity (Silver Spring)*. 2018;26(11):1785–1795. DOI: 10.1002/oby.22210.

Linge J, Cariou B, Neeland IJ, Petersson M, Rodríguez Á, Dahlqvist Leinhard O. Skewness in Body Fat Distribution Pattern Links to Specific Cardiometabolic Disease Risk Profiles. *J Clin Endocrinol Metab*. 2024;109(3):783–791. DOI: 10.1210/clinem/dgad570.

Litjens G, Kooi T, Bejnordi BE, et al. A Survey on Deep Learning in Medical Image Analysis. *Med Image Anal*. 2017;42:60–88. DOI: 10.1016/j.media.2017.07.005.

Littlejohns TJ, Holliday J, Gibson LM, et al. The UK Biobank Imaging Enhancement of 100,000 Participants: Rationale, Data Collection, Management and Future Directions. *Nat Commun*. 2020;11(1):2624. DOI: 10.1038/s41467-020-15948-9.

Liu T, Melkus G, Ramsay T, Sheikh A, Laneuville O, Trudel G. Bone Marrow Reconversion With Reambulation: A Prospective Clinical Trial. *Invest Radiol*. 2021;56(4):215–223. DOI: 10.1097/RLI.0000000000000730.

Ma J. Dixon Techniques for Water and Fat Imaging. *J. Magn. Reson. Imaging*. 2008;28(3):543–558. DOI: 10.1002/jmri.21492.

Ma Q, Cheng X, Hou X, Yang Z, Ma D, Wang Z. Bone Marrow Fat Measured by a Chemical Shift-Encoded Sequence (IDEAL-IQ) in Patients With and Without Metabolic Syndrome. *J Magn Reson Imaging*. 2021;54(1):146–153. DOI: 10.1002/jmri.27548.

Machann J, Pereira PL, Einsele H, Kanz L, Claussen CD, Schick F. [The MR characterization of the composition of the hematopoietic bone marrow. The findings in generalized neoplasms and the monitoring of therapy]. *Radiologe*. 2000;40(8):700–709. DOI: 10.1007/s001170050799.

Machann J, Thamer C, Schnoedt B, et al. Age and Gender Related Effects on Adipose Tissue Compartments of Subjects with Increased Risk for Type 2 Diabetes: A Whole Body MRI / MRS Study. *MAGMA*. 2005;18(3):128–137. DOI: 10.1007/s10334-005-0104-x.

Machann J, Hasenbalg M, Dienes J, et al. Short-Term Variability of Proton Density Fat Fraction in Pancreas and Liver Assessed by Multiecho Chemical-Shift Encoding-Based MRI at 3 T. *J Magn Reson Imaging*. 2022. DOI: 10.1002/jmri.28084.

Machann J, Horstmann A, Born M, Hesse S, Hirsch FW. Diagnostic Imaging in Obesity. *Best Practice & Research Clinical Endocrinology & Metabolism*.

Complications of Obesity 2013;27(2):261–277. DOI: 10.1016/j.beem.2013.02.003.

Machann J, Stefan N, Schick F. (1)H MR Spectroscopy of Skeletal Muscle, Liver and Bone Marrow. *Eur J Radiol*. 2008;67(2):275–284. DOI: 10.1016/j.ejrad.2008.02.032.

Machann J, Stefan N, Wagner R, et al. Normalized Indices Derived from Visceral Adipose Mass Assessed by Magnetic Resonance Imaging and Their Correlation with Markers for Insulin Resistance and Prediabetes. *Nutrients*. 2020;12(7):2064. DOI: 10.3390/nu12072064.

Machann J, Thamer C, Schnoedt B, et al. Standardized Assessment of Whole Body Adipose Tissue Topography by MRI. *J Magn Reson Imaging*. 2005;21(4):455–462. DOI: 10.1002/jmri.20292.

Machann J, Thamer C, Schnoedt B, et al. Hepatic Lipid Accumulation in Healthy Subjects: A Comparative Study Using Spectral Fat-Selective MRI and Volume-localized1H-MR Spectroscopy. *Magn. Reson. Med.* 2006;55(4):913–917. DOI: 10.1002/mrm.20825.

Machann J, Thamer C, Stefan N, et al. Follow-up Whole-Body Assessment of Adipose Tissue Compartments during a Lifestyle Intervention in a Large Cohort at Increased Risk for Type 2 Diabetes. *Radiology*. 2010;257(2):353–363. DOI: 10.1148/radiol.10092284.

Modic MT, Pflanze W, Feiglin DH, Belhobek G. Magnetic Resonance Imaging of Musculoskeletal Infections. *Radiol Clin North Am*. 1986;24(2):247–258.

Moor M, Banerjee O, Abad ZSH, et al. Foundation Models for Generalist Medical Artificial Intelligence. *Nature*. 2023;616(7956):259–265. DOI: 10.1038/s41586-023-05881-4.

Müller HP, Raudies F, Unrath A, Neumann H, Ludolph AC, Kassubek J. Quantification of Human Body Fat Tissue Percentage by MRI. *NMR Biomed.* 2011;24(1):17–24. DOI: 10.1002/nbm.1549. (Visited on 06/05/2024).

Neamat-Allah J, Wald D, Hüsing A, et al. Validation of Anthropometric Indices of Adiposity against Whole-Body Magnetic Resonance Imaging – A Study within the German European Prospective Investigation into Cancer and Nutrition (EPIC) Cohorts. *PLOS ONE*. 2014;9(3):e91586. DOI: 10.1371/journal.pone.0091586.

Neeland IJ, Ross R, Després JP, et al. Visceral and Ectopic Fat, Atherosclerosis, and Cardiometabolic Disease: A Position Statement. *Lancet Diabetes Endocrinol*. 2019;7(9):715–725. DOI: 10.1016/S2213-8587(19)30084-1.

Niedermayer F, Schauberger G, Rathmann W, et al. Clusters of Longitudinal Risk Profile Trajectories Are Associated with Cardiometabolic Diseases: Results from the Population-Based KORA Cohort. *PLoS One*. 2024;19(3):e0300966. DOI: 10.1371/journal.pone.0300966.

Ohlson LO, Larsson B, Svärdsudd K, et al. The Influence of Body Fat Distribution on the Incidence of Diabetes Mellitus. 13.5 Years of Follow-up of the Participants in the Study of Men Born in 1913. *Diabetes*. 1985;34(10):1055–1058. DOI: 10.2337/diab.34.10.1055.

OpenAI, Achiam J, Adler S, et al. *GPT-4 Technical Report*. 2024. DOI: 10.48550/arXiv.2303.08774.

Ouwerkerk R, Hamimi A, Matta J, et al. Proton MR Spectroscopy Measurements of White and Brown Adipose Tissue in Healthy Humans: Relaxation Parameters and Unsaturated Fatty Acids. *Radiology*. 2021;299(2):396–406. DOI: 10.1148/radiol.2021202676.

Pachón-Peña G, Bredella MA. Bone Marrow Adipose Tissue in Metabolic Health. *Trends in Endocrinology & Metabolism*. 2022;33(6):401–408. DOI: 10.1016/j.tem.2022.03.003.

Park CC, Hooker C, Hooker JC, et al. Assessment of a High-SNR Chemical-Shift-Encoded MRI with Complex Reconstruction for Proton Density Fat Fraction (PDFF) Estimation Overall and in the Low-Fat Range. *J Magn Reson Imaging*. 2019;49(1):229–238. DOI: 10.1002/jmri.26168.

Patzelt L, Junker D, Syväri J, et al. MRI-Determined Psoas Muscle Fat Infiltration Correlates with Severity of Weight Loss during Cancer Cachexia. *Cancers (Basel)*. 2021;13(17):4433. DOI: 10.3390/cancers13174433.

Pereira PL, Schick F, Einsele H, et al. [MR tomography of the bone marrow changes after high-dosage chemotherapy and autologous peripheral stem-cell transplantation]. *Rofo*. 1999;170(3):251–257. DOI: 10.1055/s-2007-1011036.

Peters A, German National Cohort (NAKO) Consortium, Peters A, et al. Framework and Baseline Examination of the German National Cohort (NAKO). *Eur J Epidemiol*. 2022;37(10):1107–1124. DOI: 10.1007/s10654-022-00890-5.

Peterson P, Trinh L, Månsson S. Quantitative 1H MRI and MRS of Fatty Acid Composition. *Magn. Reson. Med.* 2021;85(1):49–67. DOI: 10.1002/mrm.28471. (Visited on 07/24/2024).

Polyzotis N, Zaharia M. *What Can Data-Centric AI Learn from Data and ML Engineering?* 2021. DOI: 10.48550/arXiv.2112.06439.

Rado SD, Lorbeer R, Gatidis S, et al. MRI-based Assessment and Characterization of Epicardial and Paracardial Fat Depots in the Context of

Impaired Glucose Metabolism and Subclinical Left-Ventricular Alterations. *Br J Radiol*. 2019;92(1096):20180562. DOI: 10.1259/bjr.20180562.

Reaven GM. Banting Lecture 1988. Role of Insulin Resistance in Human Disease. *Diabetes*. 1988;37(12):1595–1607. DOI: 10.2337/diab.37.12.1595.

Reeder SB, Robson PM, Yu H, et al. Quantification of Hepatic Steatosis with MRI: The Effects of Accurate Fat Spectral Modeling. *J Magn Reson Imaging*. 2009;29(6):1332–1339. DOI: 10.1002/jmri.21751.

Reeder SB, Sirlin CB. Quantification of Liver Fat with Magnetic Resonance Imaging. *Magnetic Resonance Imaging Clinics of North America*. MRI of the Liver 2010;18(3):337–357. DOI: 10.1016/j.mric.2010.08.013.

Reinke A, Tizabi MD, Baumgartner M, et al. Understanding Metric-Related Pitfalls in Image Analysis Validation. *Nat Methods*. 2024;21(2):182–194. DOI: 10.1038/s41592-023-02150-0.

Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Med. Image Comput. Comput.-Assist. Interv. – MICCAI 2015*. Ed. by N Navab, J Hornegger, WM Wells, AF Frangi. Cham: Springer International Publishing, 2015: 234–241. DOI: 10.1007/978-3-319-24574-4_28.

Rospleszcz S, Lorbeer R, Storz C, et al. Association of Longitudinal Risk Profile Trajectory Clusters with Adipose Tissue Depots Measured by Magnetic Resonance Imaging. *Sci Rep*. 2019;9(1):16972. DOI: 10.1038/s41598-019-53546-y.

Ross R, Léger L, Morris D, de Guise J, Guardo R. Quantification of Adipose Tissue by MRI: Relationship with Anthropometric Variables. *J Appl Physiol (1985)*. 1992;72(2):787–795. DOI: 10.1152/jappl.1992.72.2.787.

Ross R, Shaw KD, Martel Y, de Guise J, Avruch L. Adipose Tissue Distribution Measured by Magnetic Resonance Imaging in Obese Women. *Am J Clin Nutr*. 1993;57(4):470–475. DOI: 10.1093/ajcn/57.4.470.

Roy AG, Conjeti S, Navab N, Wachinger C. *Inherent Brain Segmentation Quality Control from Fully ConvNet Monte Carlo Sampling*. 2018. DOI: 10.48550/arXiv.1804.07046.

Ruschke S, Syväri J, Dieckmeyer M, et al. Physiological Variation of the Vertebral Bone Marrow Water T2 Relaxation Time. *NMR Biomed*. 2021;34(2):e4439. DOI: 10.1002/nbm.4439.

Sandforth A, von Schwartzenberg RJ, Arreola EV, et al. Mechanisms of Weight Loss-Induced Remission in People with Prediabetes: A Post-Hoc Analysis of the Randomised, Controlled, Multicentre Prediabetes Lifestyle Intervention

Study (PLIS). *Lancet Diabetes Endocrinol*. 2023;11(11):798–810. DOI: 10.1016/S2213-8587(23)00235-8.

Schick F, Bongers H, Jung WI, Skalej M, Lutz O, Claussen CD. Volume-Selective Proton MRS in Vertebral Bodies. *Magn Reson Med*. 1992;26(2):207–217. DOI: 10.1002/mrm.1910260203.

Schick F, Einsele H, Weiß B, et al. Assessment of the Composition of Bone Marrow Prior to and Following Autologous BMT and PBSCT by Magnetic Resonance. *Ann Hematol*. 1996;72(6):361–370. DOI: 10.1007/s002770050187.

Schlecht I, Fischer B, Behrens G, Leitzmann MF. Relations of Visceral and Abdominal Subcutaneous Adipose Tissue, Body Mass Index, and Waist Circumference to Serum Concentrations of Parameters of Chronic Inflammation. *Obes Facts*. 2016;9(3):144–157. DOI: 10.1159/000443691.

Schlecht I, Wiggermann P, Behrens G, et al. Reproducibility and Validity of Ultrasound for the Measurement of Visceral and Subcutaneous Adipose Tissues. *Metabolism*. 2014;63(12):1512–1519. DOI: 10.1016/j.metabol.2014.07.012.

Schlett CL, Hendel T, Weckbach S, et al. Population-Based Imaging and Radiomics: Rationale and Perspective of the German National Cohort MRI Study. *Rofo*. 2016;188(7):652–661. DOI: 10.1055/s-0042-104510.

Schlett CL, Lorbeer R, Arndt C, et al. Association between Abdominal Adiposity and Subclinical Measures of Left-Ventricular Remodeling in Diabetics, Prediabetics and Normal Controls without History of Cardiovascular Disease as Measured by Magnetic Resonance Imaging: Results from the KORA-FF4 Study. *Cardiovasc Diabetol*. 2018;17:88. DOI: 10.1186/s12933-018-0721-0.

Schmeel FC, Vomweg T, Träber F, et al. Proton Density Fat Fraction MRI of Vertebral Bone Marrow: Accuracy, Repeatability, and Reproducibility among Readers, Field Strengths, and Imaging Platforms. *J Magn Reson Imaging*. 2019;50(6):1762–1772. DOI: 10.1002/jmri.26748.

Schneider M, Janas G, Lugauer F, et al. Accurate Fatty Acid Composition Estimation of Adipose Tissue in the Abdomen Based on Bipolar Multi-Echo MRI. *Magn Reson Med*. 2019;81(4):2330–2346. DOI: 10.1002/mrm.27557.

Schön M, Prystupa K, Mori T, et al. Analysis of Type 2 Diabetes Heterogeneity with a Tree-like Representation: Insights from the Prospective German Diabetes Study and the LURIC Cohort. *Lancet Diabetes Endocrinol*. 2024;12(2):119–131. DOI: 10.1016/S2213-8587(23)00329-7.

Shen W, Wang Z, Punyanita M, et al. Adipose Tissue Quantification by Imaging Methods: A Proposed Classification. *Obes. Res.* 2003;11(1):5–16. DOI: 10.1038/oby.2003.3.

Shu JB, Kim TY. Bone Marrow Adiposity in Diabetes and Clinical Interventions. *Curr Opin Endocrinol Diabetes Obes*. 2022;29(4):303–309. DOI: 10.1097/MED.0000000000000741.

Shuman WP, Morris LL, Leonetti DL, et al. Abnormal Body Fat Distribution Detected by Computed Tomography in Diabetic Men. *Invest Radiol*. 1986;21(6):483–487. DOI: 10.1097/00004424-198606000-00007.

Sneed NM, Morrison SA. Body Composition Methods in Adults with Type 2 Diabetes or at Risk for T2D: A Clinical Review. *Curr Diab Rep*. 2021;21(5):14. DOI: 10.1007/s11892-021-01381-9.

Snel M, Jonker JT, Schoones J, et al. Ectopic Fat and Insulin Resistance: Pathophysiology and Effect of Diet and Lifestyle Interventions. *Int J Endocrinol*. 2012;2012:983814. DOI: 10.1155/2012/983814.

Sollmann N, Dieckmeyer M, Schlaeger S, et al. Associations Between Lumbar Vertebral Bone Marrow and Paraspinal Muscle Fat Compositions-An Investigation by Chemical Shift Encoding-Based Water-Fat MRI. *Front Endocrinol (Lausanne)*. 2018;9:563. DOI: 10.3389/fendo.2018.00563.

Sollmann N, Kirschke JS, Kronthaler S, et al. Imaging of the Osteoporotic Spine - Quantitative Approaches in Diagnostics and for the Prediction of the Individual Fracture Risk. *Rofo*. 2022;194(10):1088–1099. DOI: 10.1055/a-1770-4626.

Sollmann N, Löffler MT, Kronthaler S, et al. MRI-Based Quantitative Osteoporosis Imaging at the Spine and Femur. *J Magn Reson Imaging*. 2021;54(1):12–35. DOI: 10.1002/jmri.27260.

Sparrow D, Borkan GA, Gerzof SG, Wisniewski C, Silbert CK. Relationship of Fat Distribution to Glucose Tolerance. Results of Computed Tomography in Male Participants of the Normative Aging Study. *Diabetes*. 1986;35(4):411–415. DOI: 10.2337/diab.35.4.411.

Starekova J, Hernando D, Pickhardt PJ, Reeder SB. Quantification of Liver Fat Content with CT and MRI: State of the Art. *Radiology*. 2021;301(2):250–262. DOI: 10.1148/radiol.2021204288.

Staten MA, Totty WG, Kohrt WM. Measurement of Fat Distribution by Magnetic Resonance Imaging. *Invest Radiol*. 1989;24(5):345–349. DOI: 10.1097/00004424-198905000-00002.

Stefan N. Identification and Characterization of Metabolically Benign Obesity in Humans. *Arch Intern Med*. 2008;168(15):1609. DOI: 10.1001/archinte.168.15.1609.

– Causes, Consequences, and Treatment of Metabolically Unhealthy Fat Distribution. *The Lancet Diabetes & Endocrinology*. 2020;8(7):616–627. DOI: 10.1016/S2213-8587(20)30110-8.

Stefan N, Fritsche A, Schick F, Häring HU. Phenotypes of Prediabetes and Stratification of Cardiometabolic Risk. *The Lancet Diabetes & Endocrinology*. 2016;4(9):789–798. DOI: 10.1016/S2213-8587(16)00082-6.

Stefan N, Schiborn C, Machann J, Birkenfeld AL, Schulze MB. Impact of Higher BMI on Cardiometabolic Risk: Does Height Matter? *The Lancet Diabetes & Endocrinology*. 2024;12(8):514–515. DOI: 10.1016/S2213-8587(24)00164-5. (Visited on 07/30/2024).

Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, Pike GB. On the Accuracy of T1 Mapping: Searching for Common Ground. *Magn. Reson. Med.* 2015;73(2):514–522. DOI: 10.1002/mrm.25135.

Suchacki KJ, Tavares AAS, Mattiucci D, et al. Bone Marrow Adipose Tissue Is a Unique Adipose Subtype with Distinct Roles in Glucose Homeostasis. *Nat Commun*. 2020;11(1):3097. DOI: 10.1038/s41467-020-16878-2.

Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med*. 2015;12(3):e1001779. DOI: 10.1371/journal.pmed.1001779.

Svendsen OL, Hassager C, Bergmann I, Christiansen C. Measurement of Abdominal and Intra-Abdominal Fat in Postmenopausal Women by Dual Energy X-ray Absorptiometry and Anthropometry: Comparison with Computerized Tomography. *Int J Obes Relat Metab Disord*. 1993;17(1):45–51.

Talman AH, Psaltis PJ, Cameron JD, Meredith IT, Seneviratne SK, Wong DTL. Epicardial Adipose Tissue: Far More than a Fat Depot. *Cardiovasc. Diagn. Ther.* 2014;4(6):41629–41429. DOI: 10.3978/j.issn.2223-3652.2014.11.05.

Tchernof A, Després JP. Pathophysiology of Human Visceral Obesity: An Update. *Physiol Rev*. 2013;93(1):359–404. DOI: 10.1152/physrev.00033.2011.

Thomas EL, Parkinson JR, Frost GS, et al. The Missing Risk: MRI and MRS Phenotyping of Abdominal Adiposity and Ectopic Fat. *Obesity*. 2012;20(1):76–87. DOI: 10.1038/oby.2011.142.

Thomas EL, Saeed N, Hajnal JV, et al. Magnetic Resonance Imaging of Total Body Fat. *J. Appl. Physiol.* 1998;85(5):1778–1785. DOI: 10.1152/jappl.1998.85.5.1778.

Tokunaga K, Matsuzawa Y, Ishikawa K, Tarui S. A Novel Technique for the Determination of Body Fat by Computed Tomography. *Int J Obes*. 1983;7(5):437–445.

Tushuizen ME, Bunck MC, Pouwels PJ, et al. Pancreatic Fat Content and Beta-Cell Function in Men with and without Type 2 Diabetes. *Diabetes Care*. 2007;30(11):2916–2921. DOI: 10.2337/dc07-0326.

Vague J. The Degree of Masculine Differentiation of Obesities: A Factor Determining Predisposition To Diabetes, Atherosclerosis, Gout, And Uric Calculous Disease. *Am J Clin Nutr*. 1956;4(1):20–34. DOI: 10.1093/ajcn/4.1.20.

Vishvanath L, Gupta RK. Contribution of Adipogenesis to Healthy Adipose Tissue Expansion in Obesity. *J Clin Invest*. 2019;129(10):4022–4031. DOI: 10.1172/JCI129191.

Völzke H, Alte D, Schmidt CO, et al. Cohort Profile: The Study of Health in Pomerania. *Int J Epidemiol*. 2011;40(2):294–307. DOI: 10.1093/ije/dyp394.

Wagner R, Eckstein SS, Yamazaki H, et al. Metabolic Implications of Pancreatic Fat Accumulation. *Nat Rev Endocrinol*. 2022;18(1):43–54. DOI: 10.1038/s41574-021-00573-3.

Wagner R, Heni M, Tabák AG, et al. Pathophysiology-Based Subphenotyping of Individuals at Elevated Risk for Type 2 Diabetes. *Nat Med*. 2021;27(1):49–57. DOI: 10.1038/s41591-020-1116-9.

Wajchenberg BL, Nery M, Cunha MR, Silva MER da. Adipose Tissue at the Crossroads in the Development of the Metabolic Syndrome, Inflammation and Atherosclerosis. *Arq Bras Endocrinol Metabol*. 2009;53(2):145–150. DOI: 10.1590/s0004-27302009000200005.

Wald D, Teucher B, Dinkel J, et al. Automatic Quantification of Subcutaneous and Visceral Adipose Tissue from Whole-Body Magnetic Resonance Images Suitable for Large Cohort Studies. *J Magn Reson Imaging*. 2012;36(6):1421–1434. DOI: 10.1002/jmri.23775.

Wang H, Leng Y, Gong Y. Bone Marrow Fat and Hematopoiesis. *Front Endocrinol (Lausanne)*. 2018;9:694. DOI: 10.3389/fendo.2018.00694.

Wang Z, Cheng C, Peng H, et al. Automatic Segmentation of Whole-Body Adipose Tissue from Magnetic Resonance Fat Fraction Images Based on

Machine Learning. *MAGMA*. 2022;35(2):193–203. DOI: 10.1007/s10334-021-00958-5.

West J, Dahlqvist Leinhard O, Romu T, et al. Feasibility of MR-Based Body Composition Analysis in Large Scale Population Studies. *PLoS One*. 2016;11(9):e0163332. DOI: 10.1371/journal.pone.0163332.

Westphalen ACA, Qayyum A, Yeh BM, et al. Liver Fat: Effect of Hepatic Iron Deposition on Evaluation with Opposed-Phase MR Imaging. *Radiology*. 2007;242(2):450–455. DOI: 10.1148/radiol.2422052024.

World Health Organization. *Obesity and Overweight*. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight. 2021. (Visited on 07/22/2022).

– *Obesity and Overweight*. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight. 2024. (Visited on 03/26/2024).

Wu B, Han W, Li Z, et al. Reproducibility of Intra- and Inter-scanner Measurements of Liver Fat Using Complex Confounder-corrected Chemical Shift Encoded MRI at 3.0 Tesla. *Sci Rep*. 2016;6:19339. DOI: 10.1038/srep19339.

Würslin C, Machann J, Rempp H, Claussen C, Yang B, Schick F. Topography Mapping of Whole Body Adipose Tissue Using A Fully Automated and Standardized Procedure. *J Magn Reson Imaging*. 2010;31(2):430–439. DOI: 10.1002/jmri.22036.

Yamazaki H, Tauchi S, Machann J, et al. Fat Distribution Patterns and Future Type 2 Diabetes. *Diabetes*. 2022;71(9):1937–1945. DOI: 10.2337/db22-0315.

Yokoo T, Serai SD, Pirasteh A, et al. Linearity, Bias, and Precision of Hepatic Proton Density Fat Fraction Measurements by Using MR Imaging: A Meta-Analysis. *Radiology*. 2018;286(2):486–498. DOI: 10.1148/radiol.2017170550.

Zeiler MD, Krishnan D, Taylor GW, Fergus R. Deconvolutional Networks. *2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 2010: 2528–2535. DOI: 10.1109/CVPR.2010.5539957.

Zelicha H, Schwarzfuchs D, Shelef I, et al. Changes of Renal Sinus Fat and Renal Parenchymal Fat during an 18-Month Randomized Weight Loss Trial. *Clin Nutr*. 2018;37(4):1145–1153. DOI: 10.1016/j.clnu.2017.04.007.

Zhan C, Olsen S, Zhang HC, Kannengiesser S, Chandarana H, Shanbhogue KP. Detection of Hepatic Steatosis and Iron Content at 3 Tesla: Comparison of Two-Point Dixon, Quantitative Multi-Echo Dixon, and MR

Spectroscopy. *Abdom Radiol (NY)*. 2019;44(9):3040–3048. DOI: 10.1007/s00261-019-02118-9.

Zhang S, Metaxas D. On the Challenges and Perspectives of Foundation Models for Medical Image Analysis. *Medical Image Analysis*. 2024;91:102996. DOI: 10.1016/j.media.2023.102996. (Visited on 07/30/2024).

Zhong X, Nickel MD, Kannengiesser SAR, Dale BM, Kiefer B, Bashir MR. Liver Fat Quantification Using a Multi-Step Adaptive Fitting Approach with Multi-Echo GRE Imaging. *Magn Reson Med*. 2014;72(5):1353–1365. DOI: 10.1002/mrm.25054.

Zhou J, Damasceno PF, Chachad R, et al. Automatic Vertebral Body Segmentation Based on Deep Learning of Dixon Images for Bone Marrow Fat Fraction Quantification. *Front Endocrinol (Lausanne)*. 2020;11:612. DOI: 10.3389/fendo.2020.00612.

Declaration of contribution of others

The work was accomplished at the Section on Experimental Radiology (Department of Diagnostic and Interventional Radiology, University Hospital Tübingen) and the Institute of Diabetes Research and Metabolic Diseases under the supervision of Prof. Dr. Jürgen Machann.
Herewith I declare that I have contributed to the major part of the publications comprising this thesis:

1. Haueise T, Schick F, Stefan N, Schlett CL, Weiss JB, Nattenmüller J, Göbel-Guéniot K, Norajitra T, Nonnenmacher T, Kauczor HU, Maier-Hein KH, Niendorf T, Pischon T, Jöckel KH, Umutlu L, Peters A, Rospleszcz S, Kröncke T, Hosten N, Völzke H, Krist L, Willich SN, Bamberg F, Machann J.
   Analysis of volume and topography of adipose tissue in the trunk: Results of MRI of 11,141 participants in the German National Cohort. *Sci Adv.* 2023;9(19):eadd0433.
   - **Tobias Haueise** selected the methods, in particular, curation of a training data set for the deep learning model, manual annotation of adipose tissue compartments, literature review and final selection of the model framework, training and evaluation of the model; performed data analysis including statistics on NAKO cohort data, in particular, processing of all data sets, automatic handling of outliers and quality control, manual inspection of outliers; interpreted the results, and wrote the manuscript.
   - Fritz Schick interpreted the results, critically reviewed the manuscript, and acquired funding for the project leading to the publication.
   - Norbert Stefan critically reviewed the manuscript.
   - Christopher L. Schlett formulated overarching research goals, was involved in the recruitment of participants, interpreted the results, critically reviewed the manuscript, and acquired funding for the project leading to the publication.
   - Johanna Nattenmüller critically reviewed the manuscript and acquired funding for the project leading to the publication.
   - Tobias Norajitra and Tobias Nonnenmacher contributed substantially to

the methods of automatic handling of outliers and quality control and critically reviewed the manuscript.

- Klaus H. Maier-Hein contributed to the methods of automatic handling of outliers and quality control, critically reviewed the manuscript and acquired funding for the project leading to the publication.
- Fabian Bamberg formulated overarching research goals, was involved in the recruitment of participants, interpreted the results, critically reviewed the manuscript, and acquired funding for the project leading to the publication.
- Jürgen Machann formulated overarching research goals and aims, selected the methods, interpreted the results, wrote and critically reviewed the manuscript, and acquired funding for the project leading to the publication.
- Jakob B. Weiss, Katharina Göbel-Guéniot, Hans-Ulrich Kauczor, Thoralf Niendorf, Tobias Pischon, Karl-Heinz Jöckel, Lale Umutlu, Annette Peters, Susanne Rospleszcz, Thomas Kröncke, Norbert Hosten, Henry Völzke, Lilian Krist, and Stefan N. Willich critically reviewed the manuscript and/or acquired the data.

2. Haueise T, Stefan N, Schulz TJ, Schick F, Birkenfeld AL, Machann J. Automated shape-independent assessment of the spatial distribution of proton density fat fraction in vertebral bone marrow. *Z Med Phys*. Published online January 30, 2023.
    - **Tobias Haueise** formulated research aims, selected the methods, in particular, curation of a training data set for the deep learning model, manual annotation of bone marrow adipose tissue, training and evaluation of the model, development of a method for shape-independent standardization of vertebral bodies; acquired the data, performed data analysis including statistics, interpreted the results, and wrote the manuscript.
    - Norbert Stefan formulated overarching research goals, recruited the participants, and critically reviewed the manuscript.
    - Tim J. Schulz formulated overarching research goals, interpreted the results, critically reviewed the manuscript, and acquired funding for the project leading to the publication.

- Fritz Schick formulated overarching research goals, interpreted the results, and critically reviewed the manuscript.
- Andreas L. Birkenfeld formulated overarching research goals, recruited the participants, and critically reviewed the manuscript.
- Jürgen Machann formulated overarching research goals and aims, selected the methods, recruited the participants, acquired the data, interpreted the results, wrote and critically reviewed the manuscript, and acquired funding for the project leading to the publication.

3. Haueise T, Schick F, Stefan N, Machann J.
   Comparison of the accuracy of commercial two-point and multi-echo Dixon MRI for quantification of fat in liver, paravertebral muscles, and vertebral bone marrow. *Eur J Radiol*. 2024;172:111359.
   - **Tobias Haueise** formulated research aims, selected the methods, in particular, curation of a training data set for the deep learning model, manual annotation of paravertebral muscles, training and evaluation of the model; acquired the data, performed data analysis including statistics, in particular, derivation of linear correction models for PDFF estimation from two-point Dixon data; interpreted the results, and wrote the manuscript.
   - Fritz Schick formulated overarching research goals, selected imaging methods, interpreted the results, and critically reviewed the manuscript.
   - Norbert Stefan formulated overarching research goals, recruited the participants, and critically reviewed the manuscript.
   - Jürgen Machann formulated overarching research goals and aims, selected imaging methods, recruited participants, acquired the data, interpreted the results, wrote and critically reviewed the manuscript, and acquired funding for the project leading to the publication.

I declare that I have written the manuscript independently and have not used any sources other than those I have indicated.

Tübingen, 9. August 2024

Tobias Haueise