Relative body weight and standardised brightness-mode ultrasound measurement of subcutaneous fat in athletes: an international multicentre reliability study, under the auspices of the IOC Medical Commission.

MÜLLER, W., FÜRHAPTER-RIEGER, A., AHAMMER, H., LOHMAN, T.G., MEYER, N.L., SARDINHA, L.B., STEWART, A.D., MAUGHAN, R.J., SUNDGOT-BORGEN, J., MÜLLER, T., HARRIS, M., KIRIHENNEDIGE, N., MAGALHAES, J.P., MELO, X., PIRSTINGER, W., REGUANT-CLOSA, A., RISOUL-SALAS, V., ACKLAND, T.R.

2019

ORIGINAL RESEARCH ARTICLE

Relative Body Weight and Standardised Brightness-Mode Ultrasound Measurement of Subcutaneous Fat in Athletes: An International Multicentre Reliability Study, Under the Auspices of the IOC Medical Commission

Wolfram Müller[1] · Alfred Fürhapter-Rieger[1] · Helmut Ahammer[1] · Timothy G. Lohman[2] · Nanna L. Meyer[3] · Luis B. Sardinha[4] · Arthur D. Stewart[5] · Ronald J. Maughan[6] · Jorunn Sundgot-Borgen[7] · Tom Müller[1] · Margaret Harris[3] · Nuwanee Kirihennedige[3] · Joao P. Magalhaes[4] · Xavier Melo[4] · Wolfram Pirstinger[1] · Alba Reguant-Closa[3] · Vanessa Risoul-Salas[3] · Timothy R. Ackland[8]

Abstract

Introduction Fat is a metabolic fuel, but excess body fat is ballast mass, and therefore, many elite athletes reduce body fat to dangerously low levels. Uncompressed subcutaneous adipose tissue (SAT) thickness measured by brightness-mode ultrasound (US) provides an estimate of body fat content.

Methods The accuracy for determining tissue borders is about 0.1–0.2 mm and reliability (experienced measurers) was within ± 1.4 mm (95% limit of agreement, LOA). We present here inter- and intra-measurer scores of three experienced US measurers from each of the centres C1 and C2, and of three novice measurers from each of the centres C3–C5. Each of the five centres measured 16 competitive adult athletes of national or international level, except for one centre where the number was 12. The following sports were included: artistic gymnastics, judo, pentathlon, power lifting, rowing, kayak, soccer, tennis, rugby, basketball, field hockey, water polo, volleyball, American football, triathlon, swimming, cycling, long-distance running, mid-distance running, hurdles, cross-country skiing, snowboarding, and ice hockey. SAT contour was detected semi-automatically: typically, 100 thicknesses of SAT at a given site (i.e., in a given image), with and without fibrous structures, were measured.

Results At SAT thickness sums D_I (of eight standardised sites) between 6.0 and 70.0 mm, the LOA of experienced measurers was 1.2 mm, and the intra-class correlation coefficient ICC was 0.998; novice measurers: 3.1 mm and 0.988. Intra-measurer differences were similar. The median D_I value of all 39 female participants was 51 mm (11% fibrous structures) compared to 17 mm (18%) in the 37 male participants.

Discussion D_I measurement accuracy and precision enables detection of fat mass changes of approximately 0.2 kg. Such reliability has not been reached with any other method. Although females' median body mass index and mass index were lower than those of males, females' median D_I was three times higher, and their percentage of fibrous structures was lower. The standardised US method provides a highly accurate and reliable tool for measuring SAT and thus changes in body fat, but training of measurers is important.

Abbreviations

BMI	Body mass index: $\mathrm{BMI} = m/h^2$ (kg m^{-2})
C	Cormic index
CT	Computer tomography (X-ray based)
$\bar{C}$	Cormic index, mean
C1–C5	Research centres
d	SAT thickness at a given site (this is the average of the distances measured within the region of interest) (mm)
D	Sum of SAT thicknesses at all eight sites in a given participant (mm)
δ	Measurement deviation of each measurement (M1, M2, and M3) from the mean of three measurements at a given site in a given subject (mm)

✉ Wolfram Müller
wolfram.mueller@medunigraz.at

Extended author information available on the last page of the article

δ_{rel}	$\delta_{rel} = 100 \cdot \delta/d_{MEAN}$ (%)
Δ	Deviation of the sum of eight sites from the mean of the sums of the three measurements in a given participant (mm)
E	Excluded; indicates that the fibrous structures are not included
ESM	Electronic supplementary material
F	Fibrous structures
F, %	Fibrous structures (%)
h	Stature (m)
I	Included; indicates that the fibrous structures are included
ISAK	International Society for the Advancement of Kinanthropometry
f	Female
l	Leg length, measured from the floor to the anterior superior iliac spine
m	Male; (also used as unit for length: metre)
m	Body mass (kg)
m_{SAT}	Subcutaneous adipose tissue mass (kg)
MI	Mass index: $MI_1 = 0.53\, m/(hs)$ (kg m^{-2})
MRI	Magnetic resonance imaging
nw	Non-weight-sensitive
ρ_{fat}	Density of body fat
s	Sitting height (m)
SAT	Subcutaneous adipose tissue
US	Ultrasound
w	Weight sensitive

Statistics

ABS	Absolute value of a number
MEAN	Mean value
N	Number of values
R^2	Coefficient of correlation
ROI	Region of interest
ρ	Spearman's rank-order correlation coefficient (Spearman's rho)
SD	Standard deviation
SEE	Standard error of estimate

US sites (order: trunk, arms, legs)

UA	Upper abdomen
LA	Lower abdomen
EO	External oblique
ES	Erector spinae
DT	Distal triceps
BR	Brachioradialis
FT	Front thigh
MC	Medial calf
LT	Lateral thigh

Key Points

Using the standardised B-mode ultrasound method, sums of subcutaneous adipose tissue (SAT) thicknesses (D) determined by experienced measurers at eight sites (on trunk, legs, and arms) can be determined with high accuracy and reliability: the 95% limit of agreement for experienced measurers (three in each of the two experienced study centres C1 and C2) was below 1.5 mm (embedded fibrous structures included: D_I), and below 2.2 mm (fibrous structures excluded: D_E). This enables monitoring changes of SAT mass in athletes (which forms the dominating part of total body fat) with an accuracy of about 0.2 kg. The median thickness measurement deviations at the individual eight sites were all below 0.2 mm. Measurement differences of novice measurers, after a 2-day course, were approximately three times larger.

This ultrasound method also allows to quantify the amount of fibrous structures (fasciae) embedded in the SAT: $D_F = D_I - D_E$. The amount of this connective tissue was significantly lower in the 39 female elite athletes of various sports (median of 11%) when compared to the 37 male elite athletes (18%). Median SAT thickness sum D_I of the eight sites was three-times higher in the elite female athletes compared to their male counterparts (51 mm vs 17 mm).

In this group of elite athletes, there was no significant correlation between SAT and body mass index (BMI). The BMI is a measure of relative body mass, but not a useful tool to determine body fat. This holds also true for the mass index MI, but this improved measure for relative body mass considers the individual's leg length, which the BMI ignores. Differences (MI-BMI) were large in several cases and ranged from -1.7 to $+1.3$ kg m^{-2} (median BMI was 22.6 kg m^{-2}), which supports the suggestion to include leg length (or sitting height) in all basic data sets of athletes and patients and thereby to assess body mass with respect to body dimensions in an improved way.

1 Introduction

In 2013, a discussion paper dealing with the question of how to minimise the health risks to athletes who compete in weight-sensitive sports was presented by the Working Group on Body Composition, Health and Performance (under the auspices of the IOC Medical Commission) [1]. This working

group also analysed advantages and shortcomings of widely used body fat assessment methods, including reference, laboratory, and field methods [2]. The authors stated: "… all of the techniques in common use have some inherent problems, whether in methodology, interpreting the data, or in the assumptions they make … Recent developments in ultrasound imaging have made possible accurate and reliable estimates of fat thickness in multiple sites of the body".

Adipose tissue layer thicknesses can be measured by a standardised ultrasound (US) approach with an accuracy not reached by any other method [2–8]. This method can be used in all persons ranging from extremely lean to obese [7]. The method has been applied in various groups, including anorectic patients [9], obese persons [7], children [10], youth athletes [11], gymnasts and swimmers [6], and rowers [11]. However, these samples were small and comprehensive data of many sports are missing. Preliminary normative data for athletes and the general population have recently been suggested [12].

The accuracy for determining tissue borders is about 0.1–0.2 mm at 12–18 MHz probe frequency [3, 6] when the appropriate speed of sound for the given tissue is used for distance determination. A detailed description of factors determining the thickness measurement accuracy can be found in the Electronic Supplementary Material (ESM). High-frequency brightness-mode US (medical diagnostic ultrasound) is the only in vivo method that enables quantifying the fibrous structures embedded in the SAT. These structures, which are not composed of adipose cells, form a substantial part of the SAT that should not be ignored when assessing "fatness". A preceding study [3] showed that the amount of these structures varied greatly, depending on both the measurement site and the person under investigation; in this studied group of 11 female football players (2nd league) and eight rhythmic gymnasts (national level), the fibrous structure median was 10%, and 50% of the values were between 6 and 17%. However, a structured analysis of the amounts of fibrous structures found in a group of male and female elite athletes of various sports has not been shown before the study presented here.

In several preceding publications [3–7], both the SAT thicknesses including the fibrous structures (indicated by the index "I") and the SAT thicknesses excluding them ("E") were measured. To measure both values is not only of interest for determining the amount of fibrous structures (i.e., the difference between these two thicknesses), but also for comparisons with other body fat measurement methods. Some techniques, such as imaging methods like magnetic resonance imaging (MRI) or computer tomography (CT), skinfolds, and cadaver studies, measure fat on the anatomical level, while others measure on the molecular level, for example the four-component model and the dual X-ray absorptiometry [2]. Accordingly,

comparisons using either SAT thicknesses including or excluding the fibrous structures are of primary relevance. Such comparative studies (using the four-component-model, DXA, and MRI high-resolution scanning) are currently in progress in the centres that contributed to this study.

Inter- and intra-measurer reliability has been tested previously [4, 6, 7], but all these comparative measurements were performed by the experienced measurers of a single centre (referred to as "C1"), and test person samples were small and not representative for a wide range of sports. In 12 elite athletes (females: five gymnasts, one swimmer; males: 4 gymnasts, 2 swimmers), with sums of SAT thicknesses D_I (from the eight standardised measurement sites) ranging from $D_I = 10$ mm to $D_I = 50$ mm, 95% of scores between measurers were within ± 1.1 mm (from the mean of the three measurements) [6]. Similar inter-measurer reliability was reached when children were studied by two experienced measurers (of the centre C1) and additionally by a third experienced measurer (of a centre that did not participate in the multicentre study described here) [13]. In a group of 38 test persons (only two elite athletes, both sailors, were in this group) ranging from lean ($D_I = 12$ mm) to obese ($D_I = 245$ mm), 95% of repeated measurements of experienced measurers of centre C1 (intra-measurer reliability) were within ± 2.2 mm. In the sub-group ranging from $D_I = 12$–77 mm, 95% of values were within ± 1.4 mm, and in a second sub-group ranging from $D_I = 44$–245 mm, 95% were within ± 2.9 mm [7]. Typical differences between measurers (ΔD_I) increased with increasing SAT thickness (d), but relative values ($\Delta D_I/d$) decreased [7].

As stated in the position statement of the IOC Working Group on Body Composition, Health and Performance [2], relative body weight determination in terms of the mass index (MI), which is a modified body mass index (BMI) considering the individual's sitting height (or leg length), should be "included in all basic data sets of athletes and patients". Considering leg length when determining relative body weight had been suggested by the Expert Committee on Physical Status of the World Health Organisation in 1995 [14].

This multicentre study aims to apply the recently developed and standardised B-mode ultrasound method for measuring SAT in a large group of elite male and female athletes ($N = 76$) of various sports to investigate the following topics:

1. Inter- and intra-measurer reliability obtained by 15 independently measuring experienced (6) and novice (9) measurers in research centres of five different countries.
2. Quantification of the SAT thickness sums and that of the fibrous structures embedded in the SAT (fasciae) in this group of elite athletes representing weight-sensitive and non-weight-sensitive sports, and to compare the

elite male athletes ($N = 37$) to the elite female athletes ($N = 39$).

3. Comparison of the SAT values of such a mixed group of elite athletes to their BMI and analysis of the impact of the individual's sitting height (which is inversely related to the leg length) on the assessment of body weight with respect to body dimensions.

2 Methods

2.1 Study Design, Participating Centres, and Measurers

SAT was measured with a recently standardised US technique [6] in five centres (C1–C5) in 36 weight-sensitive (w) and 40 non-weight-sensitive (nw) competitive athletes (national and international level). Each of the three measurers of a centre independently landmarked the eight sites on each of the athletes participating in his or her centre, and then captured and evaluated the US images. Thus, each athlete of a centre was measured three times; usually, marking, capturing the US images, and anthropometric data collection by the three measurers from each centre took place within 1 day (in the study design, the maximum time span had been fixed to 3 days). The marks on the skin were erased between the measurements of the three measurers (an erasable pen, e.g., an eyebrow pen, was used). The measurers did not have access to the results obtained by the other (two) measurers of their individual centre.

In addition, a second US image was captured in a sub-set of 56 athletes; all measurers assessed the intra-measurer participants of their centre twice: one image per site was taken, and then, this was repeated for the second image. After the first image had been captured, the US gel was removed from the skin and from the probe, and a new thick layer of gel was loaded onto the probe before the second image was taken.

Intra-measurer reliability of US image capturing and thickness evaluation was determined that way (in contrast to the inter-measurer reliability that included all three components: marking, image capturing, and thickness evaluation). Again, the measurers did not have access to the results obtained by the other measurers.

Measurers with different experiences were involved: those from centres C1–C2 were very experienced in handling the US systems (which belonged to their laboratories), and the other three centres (C3–C5) had limited experience in US imaging (and used a borrowed US system), apart from a 2-day course on US measurement and evaluation technique, followed by a supervised measurement series in five test persons.

All participants were informed about the aims and methods and gave their written consent for anonymous use of their personal data. The local ethics committees approved data collection at the five centres (in alphabetical order): Aberdeen: Robert Gordon University, UK (12-413); Colorado Springs: University of Colorado, USA (IRB14-211); Graz: Medical University of Graz, Austria (20-295 ex 08/09); Lisbon: University of Lisbon, Portugal (CEFMH, 16/2016); Perth: University of Western Australia, Australia (RA/4/1/6084). All centres declare that the study was performed in accordance with the standards of ethics outlined in the Declaration of Helsinki. The local Human Research Ethics bodies at each of the test centres follow these principles and guidelines for conducting research with human subjects.

2.2 Participants and Groups

Each of the five centres measured eight competitive adult athletes in w sports and eight in nw sports, except at C3 where the number of nw athletes was four. All athletes were in training and participated in national or international competitions. They had not undertaken strenuous exercise during the previous 48 h and reported to be normally hydrated. Inclusion criteria were as follows. Age range was 17–35 years and participants were selected from a pre-defined list of sports. Weight-sensitive sports were defined in Ackland et al. [2]; they can be summarised in three groups: aesthetic sports, weight-class sports, and gravitational sports (in which mass restricts performance due to gravitational reasons). The study design stated that all centres should capture two US images of at least eight athletes (one athlete whose weight was above his weight-class limit was eliminated; therefore, only 7 w athletes of C3 were included).

Athletes from the following sports were included: C1: w: artistic gymnastics; nw: swimming, ice hockey; C2: w: triathlon, mid-distance running, hurdles, judo, pentathlon; nw: swimming, tennis; basketball; C3: w: cross-country skiing, snowboarding, road race cycling, power lifting; nw: swimming, field hockey, soccer, rugby, open class rowing; C4: w: triathlon, mid-long-distance running; nw: swimming, water polo, kayak. C5: w: cycling, running; nw: soccer, American football, volleyball.

All 76 participants (37 males: 21 w, 16 nw; 39 females: 15 w, 24 nw) were investigated three times (by the three measurers at each centre; study part A).

For the inter-measurer analyses, athletes with $D_I > 70$ mm were excluded (exclusion of physically not well-trained athletes). This exclusion criterion resulted in a total number of 65 athletes, 30 from the expert centres C1 and C2, and 35 from the novices centres C3–C5.

A sub-group of 56 (27 w, 29 nw) participants had a second US image captured at each marked site (study part B): The five centres contributed the following numbers of athletes: C1: 8 (4 w, 4 nw); C2: 16 (8, 8), C3: 7 (1, 6), C4: 9 (6, 3), and C5: 16 (8, 8). Thus, the expert group (C1 and C2)

consisted of 24 athletes (12 w, 12 nw), and the novices group of 32 athletes (15, 17). The criterion $D_I > 70$ mm was also applied here, resulting in a total of 47 athletes participating in study part B (intra-measurer reliability).

Results obtained with all 76 athletes are presented in Fig. 2. In this context, the index "mean" refers to the means of the three measurements (three measurers measured each athlete of their centre, i.e., all athletes were measured three times). In Figs. 3 and 4, only athletes with sums of SAT thicknesses below $D_{I,mean} = 70$ mm were used; that is, only athletes with low or moderate amounts of body fat. This reduced the number of athletes of the experienced centres (C1 and C2) from 32 to 30 (i.e., 90 measurements remained). The median BMI for this sub-sample was 22.2 kg m^{-2}, ranging from 19.2 to 27.9 kg m^{-2}, and the IQR was 2.1 kg m^{-2}. The number of athlete of the three novice groups together (C3:12 athletes, C4:16, and C5:16) was reduced by this limit from 44 to 35 (i.e., 105 measurements).

The intra-measurer study (Fig. 5) included 56 participants (C1:8 athletes, C2:16, C3:7, C4:9, C5:16); the athletes of each centre were measured twice by each of the three measurers of this centre.

2.3 Anthropometry

Anthropometric measurements included body mass m, stature h, sitting height s, and leg length l (measured from the floor to the anterior superior iliac spine, ASIS). Two measures for relative body weight, the body mass index BMI $= m/h^2$ and the mass index $MI_1 = 0.53\ m/(hs)$, were calculated [15–17]. For derivation of the MI_1 equation, see ESM.

2.4 Site Marking and US Image Capture

The standardised US method has been described recently [6]. The external oblique site (EO) used in this study can cause measurement problems in obese persons and was, therefore, replaced in following measurement series by the lateral thigh site (LT), which is the thickest SAT fat depot in most women [7].

All eight US sites (Fig. 1a) were marked on the right side of the body. US images were captured with the participants lying in a defined supine, prone, or rotated position [6]. Compression of SAT was avoided using a thick layer of US gel [3, 6] between the US probe and the skin (the dark gel band can be seen on top of the US images, Fig. 1b, c). The probe was always held perpendicularly to the skin at the given site, with the centre of the probe positioned exactly above the marking.

The following brightness-mode (B-mode) US systems with linear probes operated between 12 and 18 MHz were used: C1 and C5: GE-logiq-e, probe-L8-18i, and probe 12L [General Electric, country of origin: China]; C2: Esaote Mylab One, 13 MHz [Esaote, Italy]; C3 and C4: Telemed, Echo-Blaster 128EXT-1Z-REV:C, probe-HL9.0/40/128Z-4, software Echo-WaveII-v3.2.0 [UAB Telmed, Lithuania].

2.5 Contour Detection and SAT Thickness Measurements

Ultrasound images were evaluated interactively using an evaluation software (Rotosport, Austria; rotosport.at) for semi-automatic evaluations of SAT thicknesses. Sound speed was set to 1450 m s^{-1} for distance determination in fat [3, 18–22]. At each site, the software detects the SAT segments between the lower border of the skin and the upper border of the muscle fascia (for example, see Fig. 1b, c). The algorithm measures many thicknesses (typically 50–200) within the region of interest (ROI). The mean of these thickness values is termed d and represents the SAT thickness at the site. Sums of the eight sites $D_I = d_{I,1} + \cdots + d_{I,8}$ (including embedded structures) and sums $D_E = d_{E,1} + \cdots + d_{E,8}$ (fibrous structures excluded) were calculated. Tissue segmentation was controlled visually and could be improved, if necessary, by changing the algorithm parameters. Figure 1c shows a series of eight thickness measurements for a single participant. The centre line in the US image corresponds to the central US beam; the centre of the probe was held exactly above the marking of the site. The ROI was usually set symmetrically to the centre line. A final visual control made sure that the algorithm detected the SAT layer correctly. Images in Fig. 1c correspond to the SAT patterning shown in Fig. 1d.

2.6 Statistics

SPSS (v23) software was used. Because data were not normally distributed in all sub-sets (Shapiro–Wilk test), Mann–Whitney-U test was applied for comparisons between male and female participants and between novice and experienced measurers. Correlation was tested with Spearman's rank-order correlation coefficients ρ for: D_I and BMI, D_I and MI, D_I and $D_{I,mean}$, D_E and $D_{E,mean}$, $D_{F,\%}$ and $D_{I,mean}$, $D_{I,meanB}$ and $D_{I,meanA}$, and $D_{I,AB}$ and $D_{I,meanAB}$. Limit of agreement (LOA $\approx 1.96 \cdot$ SD) [23], and intra-class correlation (according to McGraw and Wong convention) [24] were computed to quantify differences between measurers: ICC(A,1)—two-way mixed, single score, absolute agreement or ICC(A,k)—two-way mixed, average score, absolute agreement. In addition, coefficients of determination R^2 and standard errors of estimates SEE were computed. Variables and indices: D (sum of SAT thicknesses), I (fibrous structures are included), E (fibrous structures excluded), F, % (fibrous structures in %), M1 (measurer 1), M2 (measurer 2), and M3 (measurer 3). A and B denote measurement series A (inter-measurer tests) and B (intra-measurer tests).
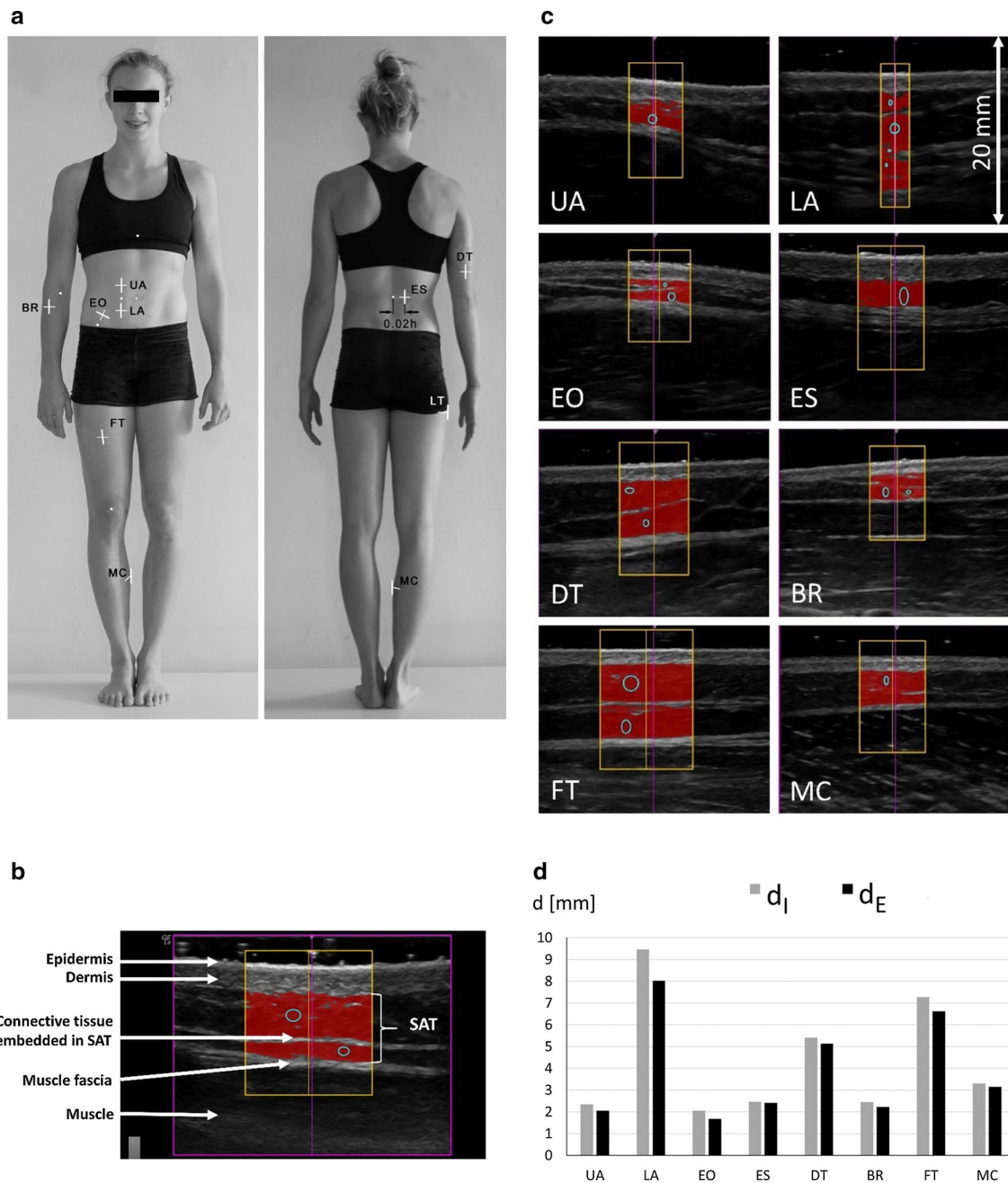
Fig. 1 Standardised ultrasound (US) measurement of subcutaneous adipose tissue (SAT). **a** US sites. *UA* upper abdomen, *LA* lower abdomen, *EO* external oblique, *ES* erector spinae, *DT* distal triceps, *BR* brachioradialis, *FT* front thigh, and *MC* medial calf. Instead of the now standard lateral thigh (LT) site, the EO site was in use when this study was performed. For a detailed description of site marking, see [6, 7]. **b** B-mode US image of SAT. All eight sites show the same structure: the black band on top represents the thick gel layer (to avoid tissue compression), followed by the structures described in the figure. The amount of embedded fibrous structures (connective tissue) varies between individuals and from site to site; this holds also true for the skin thickness. **c** Evaluated US image series. Red areas represent the SAT detected by the semi-automatic contour detection algorithm [3, 6]; the ellipses indicate the regions where the algorithm started out for SAT contour detection. In this image series, the evaluation software determined between 48 (LA) to 155 (FT) thickness values within the rectangular ROI. **d** Survey of results: SAT patterning. SAT thicknesses with the embedded fibrous structures included (d_I), and excluded (d_E). The sum of the eight d values is termed D. In this participant, D_I was 34.8 mm, D_E was 31.3 mm, and about 10% of the mean depth comprised fibrous structures, while 90% was lipid, water, and adipose cell structures. BMI of this athlete was 22.0 kg m^{-2}, and MI was 21.7 kg m^{-2}

Fig. 2 Subcutaneous adipose tissue (SAT) and embedded fibrous structures (F). In each of the five measurement centres (C1–C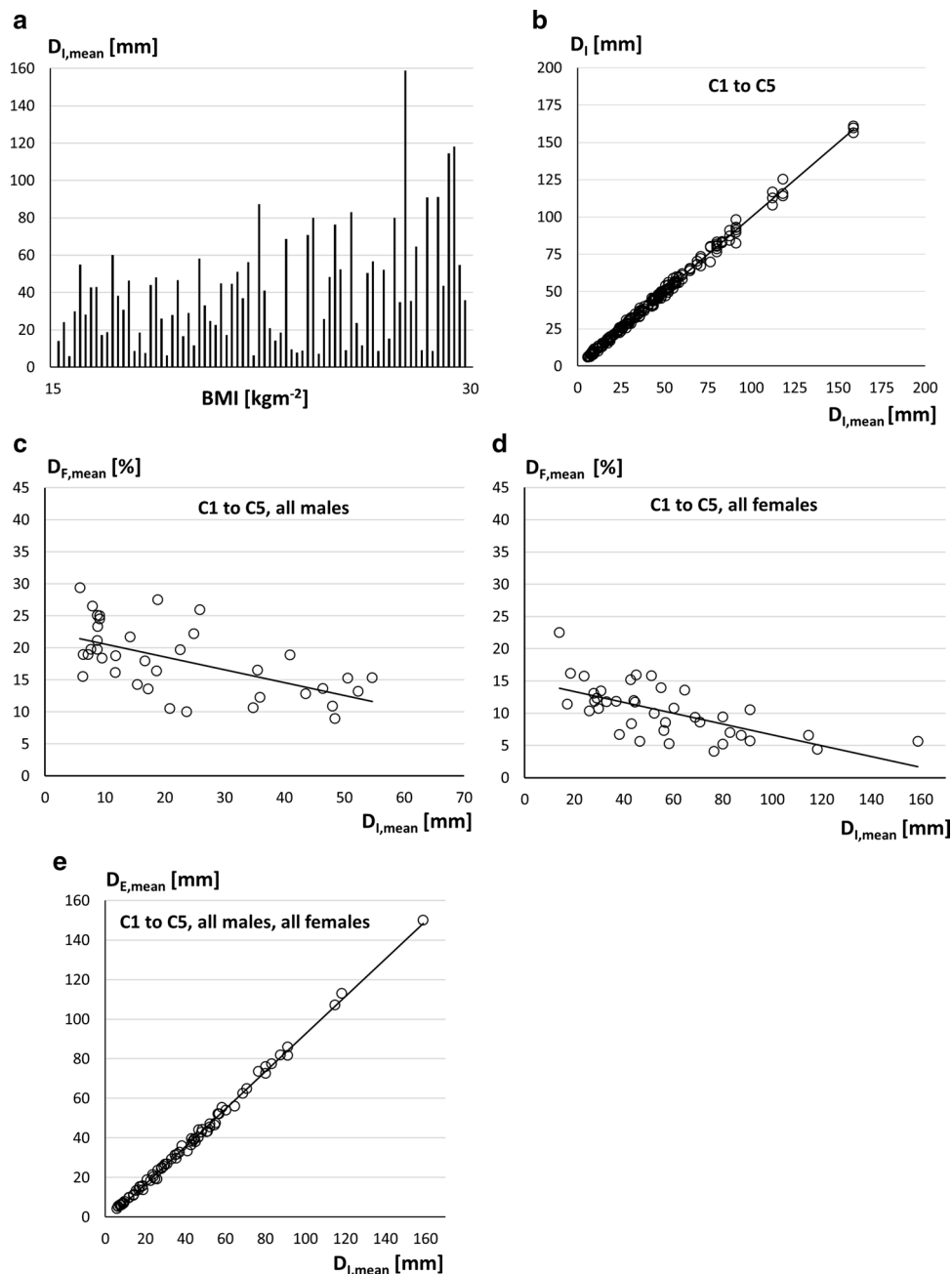5), eight weight-sensitive (w) and eight non-weight-sensitive (nw) athletes (except for C3: only four in group w) were measured three times (by the three measurers of each centre). Each of the three measurers of each centre separately marked the eight sites of all athletes of his centre, captured the according US images, and evaluated them on his own. D_I and D_E indicate sums of SAT thicknesses from the eight standardised ultrasound sites with fibrous structures included or excluded, respectively. The index "mean" indicates that means of the three measurers were used. Values of all 76 athletes (37 male and 39 female participants; 12 participants of centre C3, 16 of each other centre). The difference $D_{F,mean} = D_{I,mean} - D_{E,mean}$ represents the sum of embedded fibrous structures, and $D_{F,\%} = 100 \cdot D_{F,mean}/D_{I,mean}$. For data of individuals, see Table A1 (Electronic Supplementary Material). **a** $D_{I,mean}$ values and BMI. BMI (ranging from

17.9 to 29.0 kg m^{-2}) did not correlate with $D_{I,mean}$ ($R^2 = 0.130$, Spearman's $\rho = 0.286$). For individual values, see Electronic Supplementary Material, Table 1B. D_I ranged from 5.8 to 54.7 mm in males, and from 14.0 to 158.8 mm in females. Data of the groups are presented in Table 1 and in the Electronic Supplementary Material (Table A1 and Table A2). **b** Measurers' individual D_I results. 15 measurers (three from each centre) assessed all athletes at their centres. The coefficient of determination was $R^2 = 0.997$, $\rho = 0.997$, and SEE = 1.76 mm. **c** Percentage of embedded fibrous structures in male participants: $D_{F,\%} = 100 \cdot D_{F,mean}/D_{I,mean}$. The males' median percentage was 18.3%, $R^2 = 0.346$, and $\rho = -0.601$. **d** Percentage of embedded fibrous structures in female participants: $D_{F,\%} = 100 \cdot D_{F,mean}/D_{I,mean}$. The females' median percentage was substantially lower when compared to that of males: 10.5%, $R^2 = 0.414$, and $\rho = -0.667$. **e** Correlation of $D_{I,mean}$ and $D_{E,mean}$: $R^2 = 0.997$, and SEE = 1.5 mm

Fig. 3 Inter-measurer comparisons at experienced centres (C1–C2). Results of the three experienced measurers of each of the two centres. This sub-group of athletes, with an upper limit of $D_{I,mean} = 70$ mm (i.e., physically trained athletes with low or moderate amounts of body fat [12]), included 16 male and 14 female participants. D_I is the sum of eight SAT thicknesses d_I (fibrous structures included). D_E is the sum of eight SAT thicknesses d_E (fibrous structures excluded). **a** D_I values obtained by three experienced measurers in each athlete. ICC(A1) = 0.998, SEE = 0.6 mm, and $\rho = 0.998$. **b** D_I deviations of the three experienced measurers from their mean. SD = 0.60 mm; limit of agreement (LOA) = 1.2 mm. **c** D_E values measured by three experienced measurers in each athlete. ICC(A1) = 0.996, SEE = 0.9 mm, and $\rho = 0.996$. **d** D_E deviations of the three experienced measurers from their mean. SD = 0.9 mm; LOA = 1.7 mm
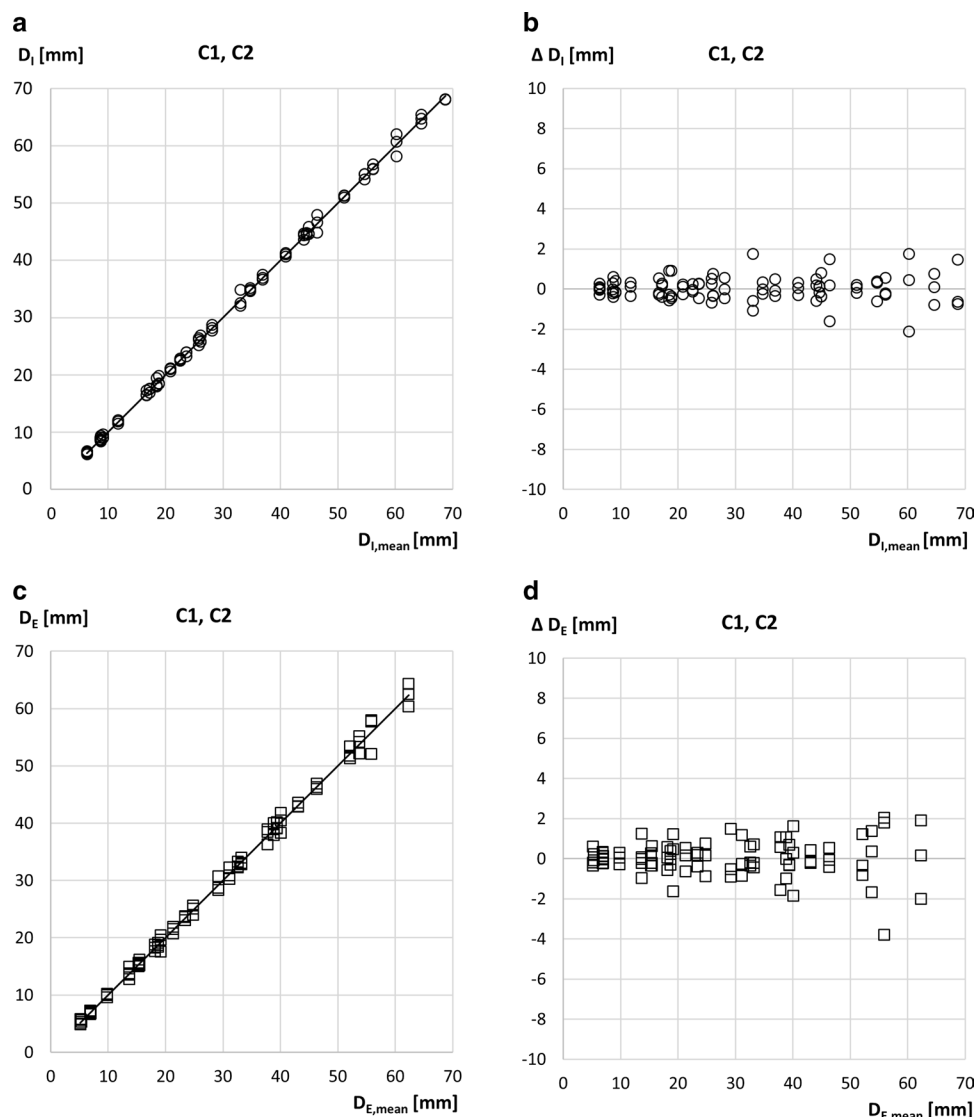
Note: Not to be mixed up with the "A" used in the ICC convention (McGraw and Wong). Box plots, medians, first and third quartiles (Q), and interquartile ranges (IQR) were used to characterise distributions.

3 Results

The survey plot (Fig. 1d) sketches both the SAT patterning with fibrous structures included in the SAT thickness (d_I) or excluded (d_E). The difference $d_F = d_I - d_E$ is the mean thickness of the embedded fibrous structures. The tissue layer thickness (d) at a given site is the mean of many (typically about 100) thickness measurements within the ROI, which is usually set symmetrically to the centre line (Fig. 1b, c). In the image series shown in Fig. 1c, the number of thickness measurements for determining the mean thicknesses d_I, d_E, and d_F at these eight sites ranged from 48 to 155.

The sum of the eight d_I values is termed D_I (sums of d_E values: D_E). Figure 1c, d shows the example of a typical female gymnast (BMI: 22.0 kg m^{-2}; MI: 21.7 kg m^{-2}): D_I was 34.8 mm, D_E was 31.3 mm, and fibrous structures $D_F = D_I - D_E$ amounted to 3.5 mm (i.e., about 10% of total SAT thickness). When this multicentre study was performed, the site EO was in use instead of LT.

In all participants ($N = 76$) of the five study centres (C1–C5), SAT thicknesses were measured three times (by the three measurers of each centre) at the eight standardised sites (Fig. 1) [6, 7]. These 1824 ultrasound (US) measurements of thickness values d, form the core data set for the inter-measurer reliability study (measurement series A). In addition, in a sub-set of 56 participants, US imaging and SAT thickness evaluation were performed twice. This second measurement series is termed B. These 1344 repeated US measurements together with the

Fig. 4 Inter-measurer comparisons at novice centres (C3–C5). This sub-group of athletes, with an upper limit of $D_{I,mean} = 70$ mm (i.e., physically trained athletes with low or moderate amounts of body fat), included 21 male and 14 female participants. D_I is the sum of eight SAT thicknesses d_I (fibrous structures included). D_E is the sum of eight SAT thicknesses d_E (fibrous structures excluded). **a** D_I values obtained by three novice measurers in each athlete. ICC(A1) = 0.988, SEE = 1.6 mm, and $\rho = 0.993$. **b** D_I deviations of the three novice measurers from their mean. SD = 1.6 mm; limit of agreement (LOA) = 3.1 mm. **c** D_E values measured by three novice measurers in each athlete. ICC(A1) = 0.977, SEE = 2.1 mm, and $\rho = 0.989$. **d** D_E deviations of the three novice measurers from their mean. SD = 2.0 mm; LOA = 4.0 mm
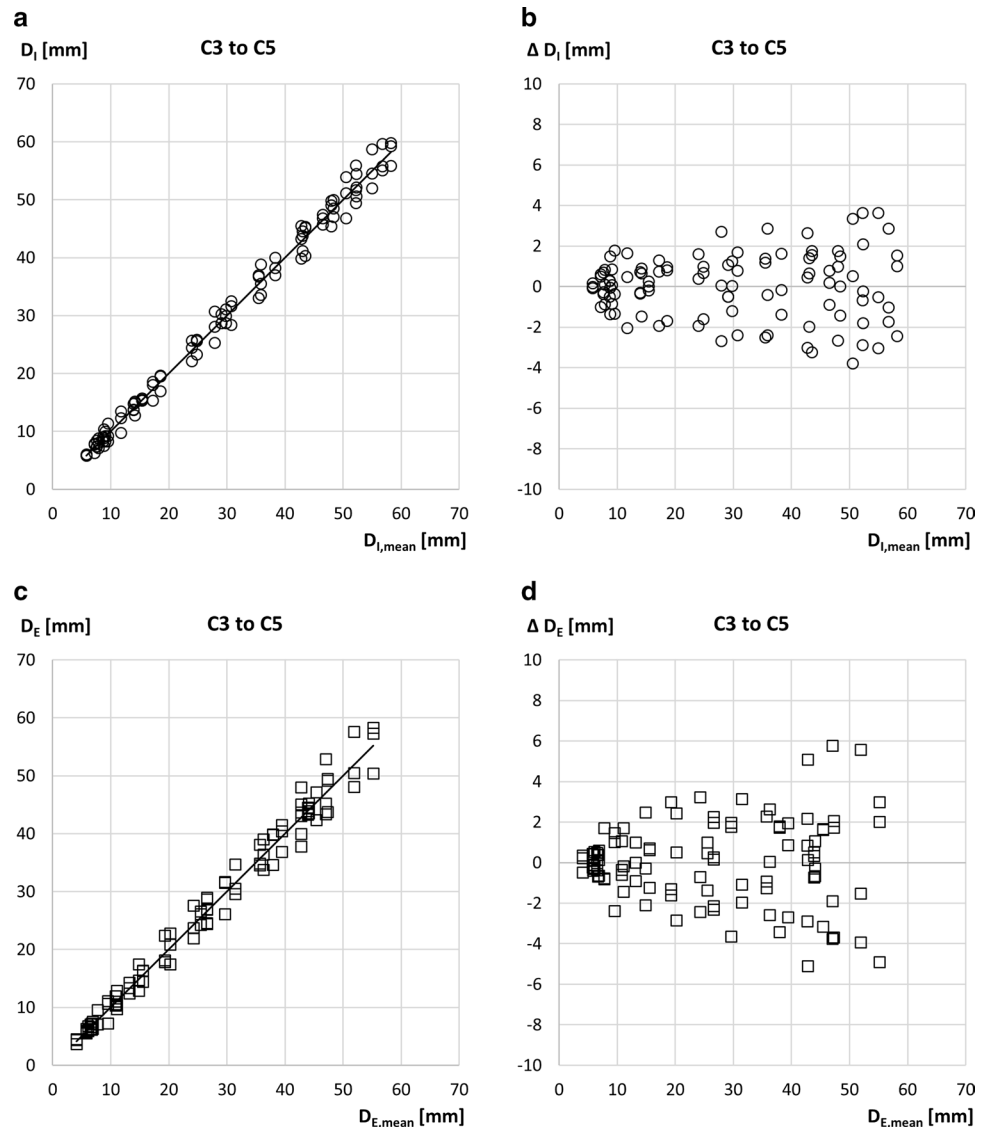
corresponding 1344 measurements of series A form the data set for the intra-measurer reliability study.

BMI was significantly lower ($p = 0.048$) according to Mann–Whitney-U in females (median was 22.1 kg m^{-2}) than in males (23.1 kg m^{-2}). This also holds true ($p < 0.01$) for the improved measure for relative body mass MI$_1$ (the MI$_1$ was not determined in centre C4, thus $N = 60$). The differences MI-BMI ranged from -1.7 to $+1.3$ kg m^{-2}. All data of individual athletes (Table A1) and of all subgroups (Table A2) are listed in the ESM.

The means of the three SAT thickness measurements of each of the 76 athletes are shown in Fig. 2a. The athletes are ordered according to increasing BMI, which ranged from 17.9 to 29.0 kg m^{-2}. There was no correlation between subcutaneous fat (D_I) and BMI ($R^2 = 0.130$, $\rho = 0.286$) or MI ($R^2 = 0.086$, SEE = 29.4 mm, and $\rho = 0.149$).

The individual results for D_I of the three measurers of all centres C1-C5 (experienced and novice measurers together) are shown in Fig. 2b: R^2 was 0.997, SEE = 1.8 mm, and $\rho = 0.997$. The respective values for D_E (not shown in a figure) were: $R^2 = 0.994$, SEE = 2.1 mm, and $\rho = 0.996$.

In the group of solely weight-sensitive (w) sports, male participants had a median of $D_I = 9.5$ mm (IQR = 16.9 mm, minimum: 5.8 mm, maximum: 55.0 mm), and female participants had a median $D_I = 33.1$ (16.3, 14.0, 55.0 mm). In the non-weight-sensitive (nw) sports group, the male participants' median was 23.1 (26.4, 8.8, 54.7 mm), and female participants' median was 66.7 (33.2, 18.5, 158.8 mm).

With regard to the embedded fibrous structures in the SAT (Fig. 2c, d), for all 76 athletes, median D_F was 4.0 mm, ranging from 1.0 to 9.5 mm, and the median percentage of embedded fibrous structures was $D_{F,\%} = 13.3\%$, ranging from 4.0 to 29.3%. For all male participants (Fig. 2c),

Fig. 5 Intra-measurer reliability comparisons. **a** A sub-group of 56 athletes was measured twice. Measurement series A and B of the three measurers are compared (Note: not to be mixed up with the "A" used in the McGraw and Wong convention for calculating the ICC values). Results of series B (mean D_I values of the three measurers) are compared to series A. Each of the series A and B scores is based on three measurements at each of the eight sites in 56 athletes (this amounts to 2688 measurements of SAT thicknesses d_I). ICC(A,k) = 0.999, SEE = 1.2 mm, limit of agreement (LOA) = 2.3 mm, and ρ = 0.999. **b** Experienced examiners only (C1–C2): a sub-group of 22 athletes from centres C1 and C2 and D_I below 70 mm. ICC(A1)$_{M1}$ = 0.999, ICC(A1)$_{M2}$ = 0.999, ICC(A1)$_{M3}$ = 0.999, ICC(A1)$_{M1-3}$ = 0.999 SEE = 0.7 mm, and ρ = 0.997. **c** Experienced examiners only (C1–C2): deviations $D_{I,B}$–$D_{I,A}$ of series A and B between the D_I measurements of the three measurers. SD = 0.7 mm; LOA = 1.4 mm. **d** Novice examiners only (C3–C5). A sub-group of 25 athletes from centres C3, C4, and C5 with D_I below 70 mm. ICC(A1)$_{M1}$ = 0.997, ICC(A1)$_{M2}$ = 0.997, ICC(A1)$_{M3}$ = 0.991, ICC(A1)$_{M1-3}$ = 0.995, SEE = 1.8 mm, and ρ = 0.988. **e** Novice examiners only (C3–C5). Deviations $D_{I,B}$–$D_{I,A}$ of series A and B between the D_I measurements of the three measurers. SD = 1.6 mm; LOA = 3.1 mm
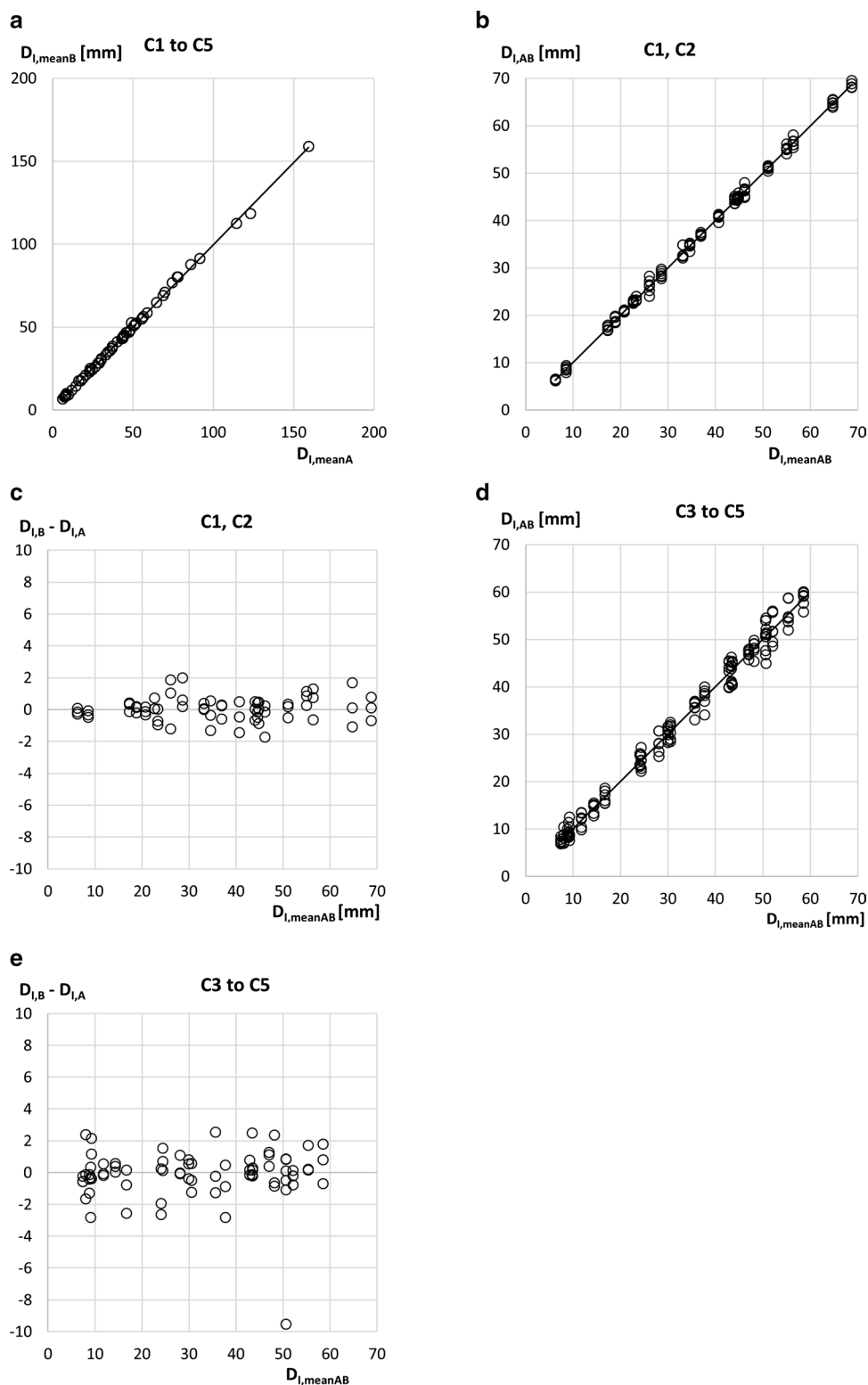
median $D_{F,\%}$ = 18.3%, ranging from 8.9 to 29.3%. In female participants (Fig. 2d), 10.5% fibrous structures were contained in the SAT, ranging from 4.0 to 22.5%. The percentage of fibrous structures tended to decrease in both men (R^2 = 0.346) and women (R^2 = 0.414) with increasing $D_{I,mean}$.

The means $D_{I,mean}$, $D_{E,mean}$, $D_{F,mean}$, and $D_{F,\%}$ differed significantly between sexes ($p < 0.001$; groups: All$_m$, All$_f$).

$D_{I,mean}$ values are closely correlated with $D_{E,mean}$ (fibrous structures excluded). Figure 2e shows the comparison of $D_{I,mean}$ and $D_{E,mean}$ for all 76 athletes (means of the three

measurers' values): the coefficient of determination R^2 was 0.997, SEE was 1.5 mm, and the slope of the regression line was 0.949. The slope was lower in the male (0.883) than in the female group (0.960), which mirrors the higher percentage of embedded fibrous structures in male athletes.

A comparison of male and female participants within the same D_I interval ranging from 14.0 to 60.0 mm (i.e., from the lowest female to the highest male value) also resulted in a significantly ($p < 0.01$) higher percentage of fibrous structures ($D_{F,\%}$) in male (median: 15%) than in female participants (12%), although the $D_{I,mean}$, $D_{E,mean}$, and $D_{F,mean}$ values did not differ significantly in these sub-groups ($p = 0.14$, $p = 0.10$, $p = 0.32$, respectively).

3.1 Inter-measurer Reliability in SAT Thickness

Inter-measurer reliability among experienced examiners (C1–C2) is compared in Fig. 3a–d. The sub-group consisted of 30 athletes (each measured three times) with an upper limit of $D_{I,mean} = 70$ mm. Median BMI for this sub-sample was 22.2 kg m^{-2}, and ranged from 19.2 to 27.9 kg m^{-2}, IQR = 2.1 kg m^{-2}.

For the D_I data of Fig. 3a, R^2 was 0.999, SEE = 0.6 mm, $\rho = 0.998$, and ICC(A,1) = 0.998. For the D_E data (Fig. 3c), R^2 was 0.997, SEE = 0.9 mm, $\rho = 0.996$, and ICC(A,1) = 0.996. Figure 3b shows the differences of the three measurers' individual D_I differences from their mean; SD = 0.6 mm, and the limit of agreement LOA = 1.2 mm. Figure 3d shows the results for D_E: SD = 0.9 mm, and LOA = 1.7 mm.

Figure 4a–d shows the comparisons of novice measurers from centres C3–C5. The sub-group consisted of 21 male and 14 female athletes (105 data points) with an upper limit of $D_{I,mean} = 70$ mm. The median BMI in this sub-group was 22.1 kg m^{-2} and ranged from 17.9 to 29.0 kg m^{-2}, IQR = 3.4 kg m^{-2}.

For the D_I data of Fig. 4a, R^2 was 0.992, SEE = 1.6 mm, $\rho = 0.993$, and ICC(A,1) = 0.988; and for the D_E data of Fig. 4c, R^2 was 0.984, SEE = 2.0 mm, $\rho = 0.989$, and ICC(A,1) = 0.977. Figure 4b shows the differences of the three measurers' individual D_I differences from their mean; SD = 1.6 mm, and LOA = 3.1 mm. Figure 4d shows the results for D_E: SD = 2.0 mm, and LOA = 4.0 mm.

Compared to experienced measurers, novices had significantly higher measurement deviations ΔD_I and ΔD_E ($p < 0.001$).

3.2 Intra-measurer Reliability in SAT Thickness

All individuals ($N = 56$) participating in the intra-measurer reliability study (measurement series B compared to the corresponding part of series A) are included in Fig. 5a. The mean scores for the three measurers obtained in series B are compared to the means of series A: $R^2 = 0.999$, SEE = 1.2 mm, LOA = 2.3 mm, $\rho = 0.999$, and ICC(A,k) = 0.999.

The results shown in Fig. 5b–e were obtained for athletes with SAT thickness sums D_I below 70 mm. The results shown in Fig. 5b, c were obtained in the sub-group of 22 athletes from centres C1–C2 (each athlete measured twice by the three measurers; median BMI = 22.4, range 19.2–27.9, IQR = 1.4 kg m^{-2}), while the results shown in Fig. 5d, e were obtained for the sub-group of 25 athletes from centres C3–C5 (median BMI = 22.0, range 18.0–26.5, IQR = 3.3 kg m^{-2}).

Figure 5b shows the measurement results of the three experienced measurers (from C1 to C2) in both series A and B (i.e., six individual measurements of the D_I values in each participant): $R^2 = 0.999$, SEE = 0.7 mm, $\rho = 0.997$, and ICC(A,1)$_{M1} = 0.999$, ICC(A,1)$_{M2} = 0.999$, ICC(A,1)$_{M3} = 0.999$, and ICC(A,1)$_{M1-3} = 0.999$. Figure 5c shows the differences in D_I between the two measurement series A and B for each of the three experienced measurers plotted against the mean scores: SD = 0.7 mm, and LOA = 1.4 mm. In addition (not shown in the figure), D_E scores were also compared: SD = 1.1 mm and LOA = 2.2 mm.

Figure 5d shows the results of the three novice measurers from each of the centres C3–C5 in both series A and B (i.e., six individual measurements of the D_I values in each participant): $R^2 = 0.989$, SEE = 1.8 mm, $\rho = 0.988$, ICC(A,1)$_{M1} = 0.997$, ICC(A,1)$_{M2} = 0.997$, ICC(A,1)$_{M3} = 0.991$, and ICC(A,1)$_{M1-3} = 0.995$. Figure 5e shows the differences in D_I between the two measurement series A and B for each of the three novice measurers plotted against the mean scores: SD = 1.6 mm, and the LOA = 3.1 mm. D_E scores: SD = 1.4 mm, LOA = 2.8 mm.

We observed a significant difference in the measurement deviations between experienced and novice measurers for the absolute values of differences $D_{I,B}-D_{I,A}$ ($p < 0.05$) and also for the absolute values of differences $D_{E,B}-D_{E,A}$ ($p < 0.01$).

For studies of fat patterning in athletes, it is of relevance to analyse the measurement differences of the three measurers at the individual sites. These differences δ (absolute values) of the three measurers from their mean are shown in the box plots of Fig. 6a–d. Measurement differences were smaller at all sites in the expert group compared to the novices for both ABS(δ_I) and ABS(δ_E). The according box plot data (Tables A3 and A4) and the

Fig. 6 Absolute values of the measurement differences ABS(δ) of the three measurers from their mean at the individual eight sites. Data are presented in the Electronic Supplementary Material (Tables A3 and A4). *UA* upper abdomen, *LA* lower abdomen, *EO* external oblique, *DT* distal triceps, *BR* brachioradialis, *LT* lateral thigh, *FT* front thigh, *MC* medial calf. Index "I": fibrous structures included, index "E": fibrous structures excluded. **a** Experienced examiners (C1, C2): ABS (δ_I) for each of the eight sites. The number of comparisons at each of the eight sites is: $N=(16+16)3=96$. **b** Novices (C3–5): ABS (δ_I) for each of the eight sites. The number of comparisons at each of the eight sites is: $N=(12+16+16)3=132$. **c** Experienced examiners: ABS (δ_E) for each of the eight sites. $N=96$. **d** Novices: ABS (δ_E) for each of the eight sites. $N=132$
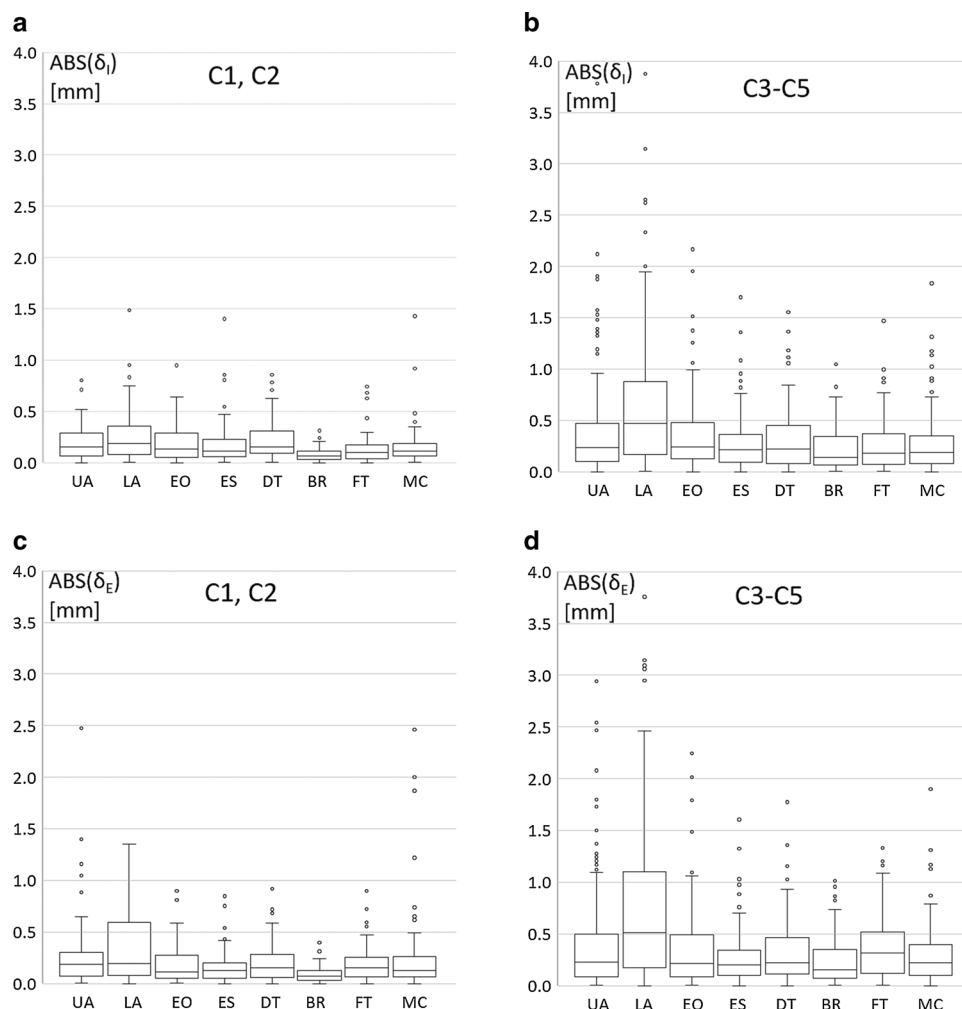
measurement differences in terms of percentages of the SAT thicknesses are presented in the ESM (Tables A5 and A6; Figs. A1a–d).

4 Discussion

The median thickness measurement deviations at the individual eight sites (inter-measurer reliability study) were all below 0.2 mm when experienced measurers performed the measurements (with sufficiently high probe frequency of about 12–18 MHz). This is comparable to the (physically given) accuracy of ultrasound distance measurement, which is mainly determined by the wavelength-dependent image resolution, provided that the correct speed of sound for the pulse-echo thickness calculation in a given tissue is used. Therefore, sums of subcutaneous adipose tissue (SAT) thicknesses can be determined with high accuracy and reliability: the 95% limit of agreement for the experienced measurers was below $D_I=1.5$ mm (and below $D_E=2.2$ mm). This enables monitoring changes of SAT mass in athletes (which

forms the dominant part of total body fat) with an accuracy of about 0.2 kg.

In female elite athletes, median SAT thickness sums D_I were three times higher as in their male counterparts (51 mm vs 17 mm). Before, only preliminary data comparing men and women have been presented [25]. B-mode ultrasound is the only imaging technique that enables also to quantify the amount of fibrous structures embedded in the SAT (fasciae). In this research, the embedded fasciae were quantified for the first time in a large group ($N=76$) of elite athletes of various sports. The amount of these connective tissues was significantly lower in the 39 female elite athletes (median: 11%) when compared to the 37 male elite athletes (18%); this further increases the ratio of subcutaneous fat in elite female athletes with respect to that in male athletes. This has not been studied before, and also comparisons of SAT amounts in male and female elite athletes have been missing; only preliminary data of a small group of non-elite athletes were available [3] and exemplary comparisons of four elite athletes [6].

For persons with the same sitting height (i.e., similar leg length), the BMI and the MI, which is a sitting height corrected BMI, are identical (definition of the MI [15–17]).

Means of BMI and MI in large groups, which represent mean sitting height can, therefore, be expected to be similar. Median BMI and MI were 22.6 kg m^{-2} and 22.2 kg m^{-2}, respectively. The small difference in our group may be because a part of the measurements were made in a Hispanic country, where sitting height medians are higher when compared to Caucasian White persons [26]; this results in MI values lower than BMI values. However, the difference between BMI and MI was large in several cases (up to 1.7 kg m^{-2}); a body mass change of more than 5 kg would correspond to a BMI difference of 1.7 kg m^{-2}. Such differences are of core relevance for both assessing the athlete's health status and for designing competition rules based on 'relative body mass' (such rules are currently used in ski jumping, for example [15, 17], where the BMI is used).

4.1 Body Fat Measurements in Sport

The status of body composition assessment in sport has been reviewed by the Working Group on Body Composition, Health and Performance (under the auspices of the IOC Medical and Scientific Commission) [2], and best practice protocols for physique assessment in sport were recently presented, including the standardised US method, which is capable of measuring SAT at an accuracy level not reached by any other method [27]. All other methods analysed there are usually not sufficiently accurate for monitoring body composition on the fine scale needed in top-level athletes. This is particularly the case if athletes are excessively small, large, or lean [2], because most athlete groups are highly specialised and their sport-specific physique imperatives are not in line with general morphological norms [27–31]. Therefore, many of the assumptions upon which measurement techniques are based are not valid in athletes. Densitometry, for example, has resulted in scores of minus 12% fat [28], and with DXA, the seven leanest in a group of male athletes showed negative fat on the torso [29]. Obviously, the morphology assumed in the measurement algorithms causes impossible results in lean athletes. Limitations of measurement techniques are discussed in the ESM and in the literature [2, 3, 6, 30–32].

4.2 Ultrasound Brightness-Mode Imaging and Distance Measurement Accuracy

Diagnostic (brightness-mode) ultrasound has been used for fat measurement since 1965 [33, 34], and many publications followed. At sufficiently high probe frequency (12–18 MHz), the thickness measurement accuracy is approximately 0.1–0.2 mm [3, 6, 35], provided that the appropriate speed of

sound in the given tissue is used (1450 ms^{-1} in fat [18–22]). The high accuracy enables measuring the embedded fibrous structures, which amount to substantial percentages of the SAT (Fig. 2c, d, Tabl.1, and ESM: Tables A1 and A2). A typical US image of SAT at the 'front thigh' site is shown in Fig. 1b. A thick layer of gel between the probe and the skin (black band above the epidermis in Fig. 1b, c) avoids compression. This is an important feature of this US measurement technique [3–7] as adipose tissue is highly compressible, and this degree of compressibility varies from site to site and between individuals [3]. Factors influencing accuracy are analysed in the ESM and in various publications [3, 6, 22]. However, the technical accuracy limits for US are not the crucial point: the limitations are set by biological reasons including detection of furrowed borders and visco-elastic deformations of adipose tissue. Therefore, measurement reliability is the overall limiting factor (Table 1).

4.3 Reliability of the Standardised US Method: An Overview

US images can never be captured by different investigators or at different times by the same investigator at exactly the same US probe position and orientation, which affects reliability. Therefore, a standardised technique has been introduced recently [6, 7]. More information about the choice of standard US sites can be found in the ESM, and in previous publications [3, 4, 6, 7]. Reliability obtained by experienced measurers has been tested in groups ranging from lean [4, 6] to overweight and obese [7]; reliability has also been tested in children [10, 13]. However, the extent to which measurer experience plays a role has never before been analysed systematically.

Tables 2 and 3 compare the core results obtained previously by experienced measurers [4, 6, 7] to the findings of the current multicentre study (MCS), in which both experienced and novice measurers were involved in the inter- and intra-measurer reliability studies. Experienced measurers of centres C1–C2 had their US system permanently available, whereas the novices (C3–C5) had to borrow a US system for their measurement series and they had no preceding experience with US imaging. Their training was limited to a 2-day course, followed by supervised US measurements in about five individuals. These are main factors causing the lower accuracy and reliability obtained by the novices.

Measurement deviations of experienced measurers in the current study (95% LOA was ± 1.2 mm for D_I that ranged from 6 to 70 mm) did not differ noticeably from previous results (± 1.0 mm, at D_I ranging from 10 to 51 mm, [6]). However, the deviations of the novice measurers were substantially larger, indicating clearly that measurers need sufficient experience to obtain the highest accuracy and

Table 1 Data of athlete groups

Groups	m (kg)	h (m)	s (m)	l (m)	MI (kg m^{-2})	BMI (kg m^{-2})	D_I (mm)	D_E (mm)	D_F (mm)	$D_{F,\%}$ (%)	N
ALL											
Median	68.2	1.731	0.931	0.953	22.2	22.6	35.1	30.4	4.0	13.3	76
IQR	14.3	0.126	0.060	0.089	2.4	2.8	35.8	32.8	3.4	8.5	
Q1	60.3	1.677	0.899	0.920	21.2	21.1	17.1	13.7	2.3	9.9	
Q3	74.6	1.802	0.959	1.009	23.6	24.0	52.9	46.5	5.8	18.4	
ALL_m											
Median	72.0	1.799	0.946	0.992	23.2	23.1	17.2	13.7	2.4	18.3	37
IQR	14.5	0.098	0.051	0.069	1.9	2.1	26.7	24.2	3.2	8.1	
Q1	66.3	1.742	0.919	0.960	22.1	22.0	8.8	6.9	2.1	13.6	
Q3	80.8	1.840	0.970	1.029	24.0	24.1	35.5	31.1	5.2	21.7	
ALL_f											
Median	61.3	1.690	0.903	0.928	21.6	22.1	51.1	44.0	4.8	10.5	39
IQR	13.8	0.068	0.056	0.069	2.1	3.4	41.7	40.7	3.0	5.8	
Q1	56.2	1.660	0.883	0.881	20.5	20.3	31.9	27.9	3.4	6.8	
Q3	69.9	1.728	0.939	0.950	22.5	23.7	73.6	68.6	6.4	12.7	
ALL_w											
Median	62.8	1.705	0.900	0.941	21.9	21.9	24.4	19.2	3.3	16.0	36
IQR	13.9	0.128	0.069	0.075	2.0	2.2	28.9	25.8	3.4	6.8	
Q1	57.0	1.662	0.872	0.913	21.2	20.8	9.0	7.0	1.9	12.8	
Q3	71.0	1.790	0.940	0.989	23.2	22.9	37.9	32.7	5.3	19.7	
ALL_nw											
Median	71.1	1.741	0.942	0.966	22.8	23.4	50.3	44.8	4.4	10.1	40
IQR	12.3	0.121	0.051	0.084	2.9	2.5	46.7	43.7	3.1	6.7	
Q1	64.9	1.698	0.917	0.928	21.1	22.0	25.5	22.9	3.0	6.9	
Q3	77.2	1.819	0.968	1.013	24.0	24.5	72.1	66.6	6.1	13.5	

Groups: *ALL* all 76 athletes, *m* males, *f* females, *w* weight-sensitive sports, and *nw* non-weight-sensitive athletes. Measured anthropometric data: *m* body mass, *h* stature, *s* sitting height, *l* leg length. Anthropometric indices: *MI* mass index, and *BMI* body mass index, *D* sum of subcutaneous adipose tissue thickness; further abbreviations: *I* fibrous structures included, *E* excluded, *F* fibrous structures, *F,%* percentage of fibrous structures with respect to d_I (thickness of the SAT at the given site), *N* number of athletes, *Q* quartile, D_I, D_E, D_F, and $D_{F,\%}$ were calculated from the mean of the three measurers' values

Data of all individuals are presented in Table A1 (Electronic Supplementary Material)

Table 2 Inter-measurer reliability

Study reference	D_I–range	95% LOA D_I	95% LOA D_E	Median of ABS (ΔD_I)	Median of ABS (ΔD_E)
2016 [6]	10–51	± 1.1	± 1.3	0.24	0.36
MCS experienced	6–70	± 1.2	± 1.7	0.30	0.38
MCS novices	6–70	± 3.1	± 4.0	0.97	1.26

Comparison of previous results [6] to the results obtained by the experienced (C1–C2) and the novice measurers (C3–C5) of this multicentre study (MCS). Results obtained in the sub-group consisting of 37 male and 28 female participants (30 of C1–C2, and 36 of C3–C5) with $D_I < 70$ mm are shown. *D* sum of SAT thicknesses (fibrous structures included: index I; excluded: index E), *LOA* limit of agreement [23]. All values in mm

reliability-level possible. The reasons for the larger errors were: bad US image quality, the US probe was not exactly at the marked position, incorrect interpretation of embedded structures as being muscle fasciae (e.g., Camper's fascia [3, 4]), the ROI not set symmetrically, or the gel layer not thick enough resulting in fat compression. Another source of error

may be that some participants did not stop breathing at mid-tidal expiration when US images were captured.

The inter- and intra-measurer deviations were larger when thicker SAT layers were measured; however, the relative deviations ($\Delta D_I/D_I$) were found to be smaller with increasing SAT thicknesses [7]. In most cases, the deviations with

Table 3 Intra-measurer reliability

Study reference	D_I–range	95% LOA D_I	95% LOA D_E	Median of ABS (ΔD_I)	Median of ABS (ΔD_E)
2017 [7]	12–77	1.4	1.6	0.43	0.41
2017 [7]	44–245	2.9	3.8	0.89	0.89
2017 [7]	12–245	2.2	3.2	0.61	0.59
MCS experienced	6–70	1.4	2.2	0.39	0.57
MCS novices	6–70	3.1	2.8	0.55	0.89

Comparison of previous results [7] to the results obtained by the experienced (C1–C2) and the novice measurers (C3–C5) of this multicentre study (MCS). Results obtained in the sub-group consisting of 25 male and 22 female participants (22 of C1–C2, and 25 of C3–C5) with $D_I < 70$ mm are shown. *D* sum of SAT thicknesses (fibrous structures included: index I; excluded: index E), *LOA* limit of agreement

All values in mm

respect to D_E (fibrous structures excluded) are slightly larger, because for measuring D_E, several tissue borders within the SAT need to be detected additionally. In the inter-measurer reliability tests, the deviations for novice measurers were about three times larger than for the experienced measurers, but in the intra-measurer reliability tests, this difference was only twofold, indicating that novices repeated some of their measurement mistakes.

4.4 Reliability at Individual Measurement Sites

The reliability of the sum D of the eight SAT thicknesses d is composed of the reliabilities of the thickness measurements at the individual sites. Figure 6a–d shows the absolute values ABS(δ) of the measurer differences from their means at the eight sites (ESM: Tables A3 and A4). Median values, interquartile ranges (IQR), and third interquartile values (Q3) were substantially smaller in the group C1–C2 (experienced examiners) compared to C3–C5 (novices) at all sites. At sites with usually higher SAT thickness d, differences ABS(δ) also tended to be higher, but all medians of the experienced group were below 0.2 mm, and below 0.5 mm in the novices' group. Not only the differences ABS(δ), in mm, but also the relative differences $\text{ABS}(\delta_{rel}) = 100 \cdot \text{ABS}(\delta)/d$, in % of the SAT thickness d at the given site, are of relevance. For example, ABS(δ) is low for EO, but the according ABS(δ_{rel}) has the highest value of all sites (ESM: Tables A5 and A6). This is one of the reasons why this site has meanwhile been replaced by lateral thigh (LT) [6]. Another reason is that the site EO causes measurement problems in obese individuals [7].

A further reason for replacing the site EO by LT is that the latter is a pronounced fat depot site in women and thus of high relevance when studying sex differences. The measurement deviations at the site LT (median of absolute deviations was 0.24 mm, median SAT thickness was 14 mm; corresponding to 1.7%) found in an intra-measurer reliability study published in 2017 [7] were comparable to the measurement deviations which these authors found at UA and LA (0.21 mm and 0.26 mm, 12 mm and 19.5 mm; 1.8%

and 1.3%, respectively). The participants studied in the cited publication [7] ranged from extremely lean to obesity class III. Based on these findings, the measurement differences at LT in our study group can be assumed to be in a similar range as found at the abdomen sites.

4.5 SAT Thickness Measurement Errors Transform Linearly into Fat Mass Errors

The small error of US thickness measurements of a fat layer transforms linearly into the error of subcutaneous fat mass, because the fat volume is proportional to the (calibrated) mean of subcutaneous fat thickness of the whole-body surface. An SAT thickness measurement error of 1.4 mm (95% LOA; see Tables 2 and 3, and Figs. 3, 4, 5) transforms into an SAT mass error of about 0.2 kg (see ESM); this is almost an order of magnitude below the daily body weight fluctuations. SAT makes by far the largest part of total body fat (typically 80–90% of anatomically detectable fat mass [36]). The SAT thickness sums in females can be expected to be higher when the site EO is replaced by LT [7, 25, 37].

None of the measurement techniques for cross-sectional or longitudinal studies of body fat is capable of measuring on such a fine scale as US [2, 27], and no other can quantify the amount of connective tissues embedded in the SAT ('fascias'), which forms a substantial part of SAT (4.0 to 29.3% in the group of elite athletes studied here).

4.6 Relative Body Mass: BMI and MI

Several indices that are power functions of body mass (m) and stature (h) were originally meant for measuring body fatness [38–40]. One such index that is widely used is the body mass index (BMI or Quetelet's index): $\text{BMI} = m/h^2$. Figure 2a shows that the BMI is useless for assessing body fat in athletes: as expected [2], there was no correlation between BMI and SAT thicknesses sums. Similar results were found in several other groups, too [7, 11, 37]. Conversely, among anorexia nervosa patients, with extremely low BMI (below

17.5 kg m^{-2}), some individuals may have subcutaneous fat thickness values comparable to those of healthy women [9, 12]. When using the BMI as a measure of 'relative body mass', there is a further important limitation that the World Health Organisation (WHO) Expert Committee on Physical Status has pointed out:

"Problems arise, however, in adults whose shape differs from the norm… Care should therefore be taken in groups and individuals with unusual leg length to avoid classifying them inappropriately as thin or overweight" [14]. Based on this justified critique, a measure for relative body mass, the mass index MI has been developed [15, 17]: $MI_1 = 0.53\, m/(hs)$. This measure considers not only stature h, but also the individual's sitting height s (and thus, implicitly, the leg length l). For the derivation of the MI_1 formula, see ESM. In this study, mean BMI was 22.6 kg m^{-2} and mean MI was 22.2 kg m^{-2}, the difference MI_1-BMI was large in several individual cases, ranging from − 1.7 to 1.3 kg m^{-2}. Particularly in weight-sensitive sports, such differences in individuals are of core relevance for assessing the athlete's health status and for rising the alarm when the individual's body weight becomes critical [1].

4.7 Characteristics of the Athlete Groups and Their SAT

Figure 2a shows that there was no correlation ($R^2 = 0.13$) between BMI (which ranged from 17.9 to 29.0 kg m^{-2}) and SAT thicknesses D_I (ranging from 6 to 160 mm). This also holds true for the MI_1 ($R^2 = 0.09$). Neither BMI nor MI_1 give useful information about athletes' body fat. Although relative body mass was 1.0 kg m^{-2} lower in females in terms of BMI (and 1.6 kg m^{-2} in terms of MI), their median D_I was 3.0-times higher (51.1/17.2 = 3.0). In addition, their median percentage of embedded fibrous structures was 1.7 times lower than in males: therefore, females' median D_E was 3.2 times the value found in males (Fig. 2c, d; Table 1). In the subgroup of athletes in weight-sensitive sports, women (median $D_I = 33.1$ mm) had about 3.5 times the amount of SAT as men (median $D_I = 9.5$ mm), and for athletes in the non-weight-sensitive group, females' median D_I (66.7 mm) was 2.9 times higher than that in males ($D_I = 23.1$). Using LT instead of EO would further increase the ratio because LT is a prominent fat depot site in women [25]. Four (of 39) D_I values of women were below 25 mm, and 15 (of 37) values of men were below 12 mm ("extremely low" according to [12]).

The means of all female participants were significantly higher for D_I, D_E, and D_F, and significantly lower for $D_{F,\%}$ when compared to means of all male participants ($p \leq 0.001$). The percentage of embedded fibrous structures tended to decrease with increasing D_I in both male and female participants ($R^2 = 0.35$ and 0.41, respectively). The median percentage of fibrous structures for all athletes was 13.3% (4.0–29.3%), for male athletes 18.3% (8.9–29.3%), and for female athletes 10.5% (4.0–22.5%).

The difference in SAT between highly trained male and female athletes is large in most cases. This also holds true for total body fat (TBF), because SAT mass represents the major part of TBF (typically 80–90%) [36].

4.8 Limitations

1. Visceral adipose tissue, which is typically about 10–20% of total body fat [36] (but may also be beyond this percentage range in some individuals), is not included in the US SAT measurement. This has to be considered when using SAT as a surrogate for total body fat.
2. Currently, only preliminary normative data are available for comparisons [12].

5 Conclusion

Regarding the reliability of this US method, when the standardised brightness-mode US technique for measuring SAT is applied by experienced measurers in athletes (with low or moderate body fat: $D_I < 70$ mm), the 95% LOA can be expected to be below 1.5 mm for the sum of thicknesses from eight sites D_I (fibrous structures included), and below 2.2 mm for D_E (fibrous structures excluded). At the individual eight sites, median measurement differences (from their means) ranged from 0.06 to 0.19 mm (third quartiles: 0.11 to 0.36 mm).

The inter-measurer results found here in a large group of athletes of various sports are in line (Table 2) with a preliminary study [6] that compared the results of three experienced measurers obtained in a small group of lean athletes ($N = 12$). The standardised US method enables tracking of SAT thickness changes that correspond to about 0.2 kg changes in SAT mass, which is substantially below the daily body mass changes. Measurement differences of novice measurers were approximately three times larger. Their results are still useful; however additional training, particularly in US image capturing, is necessary to attain the highest possible level of reliability.

In terms of body composition, the US measurement results obtained in this group of elite athletes from various sports showed that the median SAT thickness sum D_I was three times higher in the elite female athletes as in the male group (51 mm vs 17 mm). In addition, the percentage of connective tissue embedded in the SAT was significantly ($p < 0.01$) lower in women (median 11%) than in men (18%) and percentages tended to decrease with increasing D in both groups. This also holds true when comparing female (12%) and male (15%) participants within the same D_I interval ranging from the lowest female to the highest male value (14–60 mm), although the D_I, D_E, and D_F values

did not differ significantly in these sub-groups ($p > 0.1$). The standardised US method is the only measurement technique that has sufficient accuracy to quantify the amount of fibrous structures embedded in the SAT.

Comparing BMI to D_I, there were no correlations between BMIs or MIs and SAT thickness sums D_I for this group of all 76 athletes from the five research centres (D_I ranged from 6 to 160 mm, and BMI from 18 to 29 kg m^{-2}). The BMI is a measure of relative body mass, but not a useful tool to determine body fat. This also holds true for the MI, but this improved index of body mass considers the individual's leg length. Differences (MI-BMI) were large in several cases and ranged from -1.7 to $+1.3$ kg m^{-2} (median BMI was 22.6 kg m^{-2}), which supports the inclusion of sitting height (or leg length) in all basic data sets.

Comparing D_I to D_E, the values of SAT sums with fibrous structures included (D_I) are closely correlated with D_E ($N = 76$ athletes, for means of the three measurers' values: $R^2 = 0.997$, SEE $= 1.5$ mm); the slope of the regression line was lower in the male (0.88) than in the female group (0.96), indicating the higher percentage of embedded fibrous structures in male athletes.

Future research We encourage the application of this standardised US method for the study of body composition in athletes of various sports, and the use of these data sets for performance optimisation or medical diagnoses for sports in which low weight and body composition problems exist. Only preliminary data sets are currently available, so the question of what minimum fat level is acceptable from a medical point of view for an individual with unique genetics and lifestyle cannot be answered at this time. In a recent study of anorectic patients [9], the US method has shown that SAT amounts differed by 330%, although their (extremely low) BMI differed by only 12% when the 18 female patients were divided into two groups based on the group median of D_I.

The standardised US method enables accurate studies of fat patterning. In this study, we primarily discussed the comparison of sums of SAT thicknesses from the eight sites between male and female athletes, but there is much more information contained in the distribution of fat in men and women of different sport groups, and in patients with chronic conditions. Other lines of research include intervention studies that effect body composition (e.g., studies on physical training effects or sports nutrition). Furthermore, it will be interesting to see how the US data compare to other established methods for measuring body composition, like the four-component model, MRI, or DXA. Such studies are in progress in some of the centres that participated in this study.

Compliance with Ethical Standards

References

1. Sundgot-Borgen J, Ackland TR, Lohman TG, et al. How to minimise risks for athletes in weight-sensitive sports: review and position statement on behalf of the Ad Hoc Research Working Group on Body Composition Health and Performance, under the auspices of the I.O.C. Medical Commission. Br J Sports Med. 2013;47(16):1012–22.
2. Ackland T, Lohman T, Sundgot-Borgen J, et al. Current status of body composition assessment in sports. Rev Article Sports Med. 2012;42(3):227–49.
3. Müller W, Horn M, Fürhapter-Rieger A, et al. Body composition in sport: a comparison of a novel ultrasound imaging technique to measure subcutaneous fat tissue compared with skinfold measurement. Br J Sports Med. 2013;47:1028–35.
4. Müller W, Horn M, Fürhapter-Rieger A, et al. Body composition in sport: inter-observer reliability of a novel ultrasound measure of subcutaneous fat tissue. Br J Sports Med. 2013;47:1036–43.
5. Müller W, Maughan RJ. The need for a novel approach to measure body composition: Is ultra-sound an answer? Br J Sports Med. 2013;47(16):1001–2.
6. Müller W, Lohman TG, Stewart A, et al. Subcutaneous fat patterning in athletes: selection of appropriate sites and standardisation of a novel ultrasound measurement technique. Br J Sports Med. 2016;50:45–54.
7. Störchle P, Müller W, Sengeis M, et al. Standardized ultrasound measurement of subcutaneous fat patterning: high reliability and accuracy in groups ranging from lean to obese. Ultrasound Med Biol. 2017;43(2):427–38.
8. Meyer NL, Ackland TR, Lohman TG, et al. Body composition for Health and Performance: a Survey of the Ad Hoc Research Working Group on Body Composition, Health and Performance, under the auspices of the IOC Medical Commission. Br J Sports Med. 2013;47:1044–53.

9. Lackner S, Mörkl S, Müller W, et al. Novel approaches for the assessment of relative body weight and body fat in diagnosis and treatment of anorexia nervosa: a cross-sectional study. Clin Nutr. 2019. https://doi.org/10.1016/j.clnu.2018.12.031.

10. Kelso A, Vogel K, Steinacker JM. Ultrasound measurements of subcutaneous adipose tissue thickness show sexual dimorphism in children of three to five years of age. Acta Paediatr. 2018. https://doi.org/10.1111/apa.14496.

11. Kelso A, Trájer E, Machus K, et al. Assessment of subcutaneous adipose tissue using ultrasound in highly trained junior rowers. Eur J Sport Sci. 2017;17(5):576–85.

12. Ackland T, Müller W. Imaging method: ultrasound. In: Hume P, Kerr S, Ackland T, editors. Best practice protocols for physique assessment in sport. New York: Springer; 2018. p. 131–41.

13. Kelso A, Müller W, Fürhapter-Rieger A. Inter-observer reliability in standardised ultrasound measurements of subcutaneous adipose tissue in children. 2018 **(submitted)**.

14. WHO Expert Committee. Physical status, use and interpretation of anthropometry. Tech Rep Ser. 1995;854:355.

15. Müller W. Determination of ski jump performance and implications for health, safety and fairness. Sports Med. 2009;39(2):85–106.

16. Müller W. Towards research based approaches for solving body composition problems in sports: ski jumping as a heuristic example. Br J Sport Med. 2009;43(13):1013–9.

17. Müller W, Gröschl W, Müller R, et al. Underweight in ski jumping: the solution of the problem. Int J Sports Med. 2006;27(11):926–34.

18. Herman IP, editor. Physics of the human body, 1st ed. New York: Springer; 2007. p. 559.

19. Snyder WS, Cook MJ, Nasset ES, et al. Report of the task group on reference man. Commission on Radiological Protection No. 23. Oxford: Pergamon Press; 1975 **(reprinted 1992,49)**.

20. Koch T, Lakshmanan S, Brand S, et al. Ultrasound velocity and attenuation of porcine soft tissues with respect to structure and composition: II. Skin and backfat. Meat Sci. 2011;88(1):67–74.

21. Goss S, Johnston R, Dunn F. Comprehensive compilation of empirical ultrasonic properties of mammalian tissues. J Acoust Soc Am. 1978;64:442.

22. Störchle P, Müller W, Sengeis M, et al. Measurement of mean subcutaneous fat thickness: eight standardised ultrasound sites compared to 216 randomly selected sites. Sci Rep. 2018;8(1):16268.

23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurment. Lancet. 1986;1:307–10.

24. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1:30–46.

25. Störchle P, Sengeis M, Müller W. FAT patterning in males and females: application of a highly accurate and precise ultrasound technique. Vienna: ECSS Congress; 2016.

26. Norgan NG. Population differences in body composition in relation to the body mass index. Eur J Clin Nutr. 1994;48(3):10–27.

27. Hume P, Kerr S, Ackland T, editors. Best practice protocols for physique assessment in sport. 1st ed. New York: Springer; 2018.

28. Adams J, Mottola M, Bagnell KM, et al. Total body fat content in a group of professional football players. Can J Appl Sport Sci. 1982;7:36–40.

29. Stewart AD, Hannan WJ. Prediction of fat and fat free mass in male athletes using dual X-ray absorptiometry as the reference method. J Sports Sci. 2000;18:263–74.

30. Moore TL, Lunt M, McManus B, et al. Seventeen-point dermal ultrasound scoring system—a reliable measure of skin thickness in patients with systemic sclerosis. Rheumatology. 2003;42:1559–63.

31. Marfell-Jones Nevill AM, Stewart AD. Anthropometric surrogates for fatness and health. In: Stewart AD, Sutton L, editors. Body composition in sport, exercise and health. Abingdon: Routledge; 2012. p. 126–46.

32. Kerr A, Hume PA. Non-imaging method: bioelectrical impedance analysis. In: Hume P, Kerr S, Ackland T, editors. Best practice protocols for physique assessment in sport. 1st ed. New York: Springer; 2018. p. 101–16.

33. Bullen BA, Quaade F, Olsen F, et al. Ultrasonic reflections used for measuring subcutaneous fat in humans. Hum Biol. 1965;37:375–84.

34. Booth RAD, Goddard BA, Paton A. Measurement of fat thickness in man: a comparison of ultrasound, Harpenden calipers and electrical conductivity. Br J Nutr. 1966;20:719–25.

35. Bellisari A, Roche AF. Anthropometry and ultrasound. In: Heymsfield SB, Lohman TG, Wang ZM, Going SB, editors. Human body composition. 2nd ed. Champaign: Human Kinetics; 2005. p. 109–27.

36. Ibrahim MM. Subcutaneous and visceral adipose tissue: structural and functional differences. Obes Rev. 2010;11:11–8.

37. Sengeis M, Müller W, Störchle P, et al. Body weight and subcutaneous fat patterning in elite judokas. Scand J Sports Med. 2019. https://doi.org/10.1111/sms.13508.

38. Lee J, Kolonel LN, Hinds MW. Relative merits of the weight-corrected-for-height indices. Am J Clin Nutr. 1981;34:2521–9.

39. Keyes A, Fidanza F, Karvonen MJ, et al. Indices of relative weight and obesity. J Chron Dis. 1972;25:329–43.

40. Benn RT. Some mathematical properties of weight-for-height indices used as measures of adiposity. Br J Prev Soc Med. 1971;25:42–50.

Affiliations

Wolfram Müller[1] · Alfred Fürhapter-Rieger[1] · Helmut Ahammer[1] · Timothy G. Lohman[2] · Nanna L. Meyer[3] · Luis B. Sardinha[4] · Arthur D. Stewart[5] · Ronald J. Maughan[6] · Jorunn Sundgot-Borgen[7] · Tom Müller[1] · Margaret Harris[3] · Nuwanee Kirihennedige[3] · Joao P. Magalhaes[4] · Xavier Melo[4] · Wolfram Pirstinger[1] · Alba Reguant-Closa[3] · Vanessa Risoul-Salas[3] · Timothy R. Ackland[8]

1   Biophysics, Medical University of Graz, Neue Stiftingtalstraße 6, 8010 Graz, Austria

2   University of Arizona, Tucson, USA

3   University of Colorado Colorado Springs, Colorado Springs, CO, USA

4   Faculdade Motricidade Humana, CIPER, Universidade Lisboa, Lisbon, Portugal

5   Robert Gordon University, Aberdeen, UK

6   School of Medicine, St Andrews University, St Andrews, UK

7   NIH, The Norwegian School of Sport Sciences, Oslo, Norway

8   University of Western Australia, Perth, Australia

Electronic Supplementary Material (ESM)

Relative body weight and standardised ultrasound measurement of subcutaneous fat in athletes: an international multicentre reliability study, under the auspices of the IOC Medical Commission
(Sports Medicine)

Wolfram Müller, Alfred Fürhapter-Rieger, Helmut Ahammer, Timothy G. Lohman, Nanna L. Meyer, Luis B. Sardinha, Arthur D. Stewart, , Ronald J. Maughan, Jorunn Sundgot-Borgen, Tom Müller, Margaret Harris, Nuwanee Kirihennedige, Joao P. Magalhaes, Xavier Melo, Wolfram Pirstinger, Alba Reguant-Closa, Vanessa Risoul-Salas, Timothy R. Ackland

Corresponding author:
Wolfram Müller
Medical University of Graz, Biophysics
Neue Stiftingtalstraße 6
8010 Graz, Austria
Email: wolfram.mueller@medunigraz.at
Phone: ++43 316 380 3913

Electronic Supplementary Material

Ad: Obtainable accuracy of US thickness measurements

The accuracy obtainable with brightness-mode (B-mode) US depends on the probe frequency (high frequency increases resolution, but reduces image depth due to increased attenuation), on the appropriate setting of the US system (primarily: gain, time-gain compensation, image depth, number and position of foci, image analysis parameters), and on the skills of the investigator (small movements or tilts of the probe can change image quality substantially). Linear probes should be used for quantitative measurements. Tissue compression can be avoided by including a thick layer of gel between the probe and the skin [3,6]. The resolution of US imaging is determined by the US wavelength (λ): at 18 MHz probe frequency (f), which is the best choice for the thin subcutaneous adipose tissue layers found in most athletes, a resolution approximately equal to the wavelength (0.1 mm) can be obtained because diffraction and spatial length of the US pulse are the limiting factors for transverse and longitudinal resolution: $\lambda = c/f = 1450 \text{ m s}^{-1}/1.8 \cdot 10^7 \text{s}^{-1} \approx 0.1 \text{ mm}$. A sound speed deviation from the real speed in the given tissue of 30 ms^{-1} would result in 2% thickness measurement error. For example, a 10 mm thick layer would erroneously be measured as 10.2 mm. Sound velocity data can be found in [18-21] and detailed discussions of the US thickness measurement accuracy can be found in [3,6,22].

Note: It is not the technically given high accuracy that limits the usefulness of the US method, but rather biological reasons (furrowed borders and visco-elastic deformations of adipose tissue) mostly affect the reliability. The technically given accuracy of about 0.1 mm cannot be outperformed by any other method (including mechanical measurements of tissue thickness layers using a micrometre screw [2]).

Choice of US sites

Deviations of repeated measurements depend primarily on the micro-anatomical structure at a given site (provided that the measurers are trained sufficiently). In most cases, these deviations are larger than the technically given accuracy limitations. To obtain the highest possible reliability, it is of paramount importance to choose US sites where SAT thickness is mostly constant around the centre of the site, to choose sites

that can be easily and reliably marked, and to capture the US images in standardised positions (because fat is visco-elastic) [3,4,6,7]. The sites selected for the standardised US measurement of SAT obey these criteria and they represent trunk (3 sites), arms (2 sites), and legs (3 sites).

Linear transformation of SAT thickness measurement error into fat mass error

The small errors resulting from the accuracy and reliability limitations of US thickness measurements of SAT layers transform linearly into the error of subcutaneous adipose tissue mass m_{SAT} because the fat volume is proportional to the (calibrated) mean subcutaneous fat thickness ($d_{mean,cal}$): $m_{SAT} = S\, d_{mean,cal}\, \rho_{fat}$, with S being the surface area, and ρ_{fat} the density of fat tissue. As usual with skinfolds [31], we believe that the sum of thicknesses should be used directly for comparisons between athletes because any calculation of total body fat mass (or of total SAT mass) involves model assumptions and includes possible additional measurement errors (surface area, mean SAT thickness calibration, density of SAT in various body regions and individuals, changing water content). However, for a rough estimate in an athlete with a body surface area of *A=2.0 m²*, a D_I-value of 8 mm would result in about 1.4 kg SAT. For this approximate calculation, a factor of 0.7 was used for calibrating the mean SAT thickness derived from the eight sites [22] (this considers that the standardised sites over-represent the mean SAT thickness because these sites were chosen to represent typical fat patterning, but not mean SAT thickness). From this approximate assessment of fat mass, we learn that a measurement error of 1.4 mm (95% LOA; see Tables 2 and 3) transforms into a SAT mass error of less than 0.2 kg.

Body fat measurements in sport:

In all computed tomography methods, the accuracy of volumetry (and thus the determination of fat mass) depends strongly on the choice of the image segmentation parameters and the measurement parameter settings. Additionally, MRI pixel size used for total body scans (typically 1.3-2.0 mm) is not small enough for measuring the thin fat layers in lean athletes with sufficient accuracy. X-ray computer-tomography (CT) cannot be used for routine measurements because of the high radiation exposure. Skinfolds (SF) measure a compressed double layer of SAT and skin: a comparison of highly accurate US thickness measurements of SAT to SF showed that compressibility

of SAT and also skin thickness (variation of 300%) largely depends on the anatomical site and on the individual [3,30]. This explains the accuracy limitations of SF methods, which can be severely misleading [3]. Fat mass calculations for individuals based on SF is outdated because of excessive errors [31]. Similarly, large errors result from the widely used bioimpedance (BIA) method [2,8,27], which is based on a highly reductionist measurement concept that aims to assess body fat mass by simply measuring the alternating current resistance (e.g., at 50 kHz) of the human body [2,27]. For detailed discussion of methods in use, see [2,8,27].

BMI and the derivation of the MI formula for relative body weight:

The body mass index (Quetelet's index) $\mathrm{BMI} = m/h^2$ (m body mass, h stature) is not useful for assessing body fat of individuals, particularly in athletes (compare to Fig. 2a in the main text). When using the BMI for assessing 'relative body mass', there is a further important limitation: the BMI ignores individual body properties [14]. Therefore, W. Müller has developed an improved measure for relative body weight, which considers the sitting height s (and thus, implicitly, also the leg length l): the mass index MI [15,16,17]. The MI is a modified BMI. The general formula for this modified BMI is:

$$\mathrm{MI} = \mathrm{BMI} \, (\overline{\mathrm{C}}/\mathrm{C})^k \text{ , or, explicitly: } \mathrm{MI} = \frac{m}{h^2} \left(\frac{\overline{\mathrm{C}}}{\frac{s}{h}} \right)^k$$

The exponent k weights the modification term. The Cormic index $\mathrm{C} = s/h$. At a given h, a large s is associated with a low l (according to $h=s+gl$; g is a geometry factor of the individual). Therefore, both s or l can be used for designing a measure that considers the individual's geometric properties. $\overline{\mathrm{C}} = 0.53$ represents mean sitting height (this reference value is chosen in the middle of the Cormic index continuum). The choice of $k = 1$ (thus 'MI₁') takes both into consideration, the dependency of relative body weight on stature h and on sitting height s:

$$\mathrm{MI}_1 = 0.53 \cdot \frac{m}{hs} \ .$$

MI₁ is chosen symmetrically between $\mathrm{MI}_2 = 0.53^2 \cdot m/s^2$ ($k = 2$ would ignore the impact of stature h on relative body mass as h would cancel out in this case), and $k = 0$ ignores the impact of sitting height s (and thus of leg length, too):

$\mathrm{MI}_0 = m/h^2 \equiv \mathrm{BMI}$. For a person with long legs, the MI is higher than the BMI, and vice versa for a person with short legs. With the same stature h, a person with shorter legs (large sitting height) can be expected to have higher body mass m because his volume

is higher due to the relatively large dimensions of his upper body. Particularly in low weight sports, assessing 'relative body mass' of athletes is a crucial health parameter. A BMI of 17.5 kgm^{-2} is one of the four criteria for diagnosing anorexia nervosa: when using the MI instead of the BMI, different diagnoses would result whether the MI is equal to the BMI (due to mean leg length), or differs by one or even more units. In the group of elite athletes described in this publication, differences MI-BMI ranged from -1.7 kgm^{-2} to +1.3 kgm^{-2}. At a BMI of 17.5 kgm^{-2}, a difference of -1.7 kgm^{-2} would result in 15.8 kgm^{-2}, which is far below the weight criterion for anorexia nervosa, whereas +1.3 kgm^{-2} would result in 18.8 kgm^{-2}, which is far above this criterion, and even above the WHO criterion for underweight (which is 18.5 kgm^{-2}). In addition to using the MI, the accurately measured SAT amount should be included in diagnostics and therapeutics of low weight problems in athletes and in anorexia nervosa patients [9].

The MI is defined such that the WHO cut-off points for underweight, overweight, and obesity (18.5, 25, and 30 kgm^{-2}) can remain the same when replacing the BMI by the MI: their means are equal for groups with a mean Cormic Index of C = 0.53. The most important advantage of the MI over the BMI is the appropriate assessment of relative weight of the individual, although group means may be similar or the same. In the group of mainly Caucasians (and a small number of Hispanics) studied here, median BMI was 22.6 kgm^{-2}, and median MI 22.2 kgm^{-2}. Using the MI_1 instead of the BMI will also contribute to the discussion about BMI cut-off points for ethnic groups with shorter or longer legs [26]. These populations show large differences between MI and BMI.

Table A1: Individual data of all 76 participating athletes.
The same abbreviations as in Table A2 are used here. Additionally: w (weight-sensitive), and nw (non-weight-sensitive). Ordered according to increasing body mass index (BMI).

Table A2: Data of athlete sub-groups
Groups: all 76 athletes (ALL), males (m), females (f), weight sensitive sports (w), and non-weight-sensitive athletes (nw). C1 to C5 indicate the athlete sub-groups of the five participating centres. Anthropometric data: body mass (*m*), stature (*h*), sitting height (*s*), leg length (*l*). *Anthropometric* indices: mass index MI, and body mass index BMI. D: sum of subcutaneous adipose tissue thickness. Further abbreviations: mean: mean of the values of the three measurers; I: fibrous structures (i.e., fasciae) included; E: excluded; F: fibrous structures (fasciae); F,%: percentage of fibrous structures (fasciae) with respect to D_I. N is the number of athletes. Note: not all data sets are normally distributed. For median values of the main groups, see Table 1 (in the main text). Data are sorted according to increasing BMI.

Table A1

m	h	s	l	MI	BMI	MI-BMI	D_I,mean	D_E,mean	D_F,mean	D_F,%	Age	Sex		Sport		Sports
[kg]	[m]	[m]	[m]	[kgm⁻²]	[kgm⁻²]	[kgm⁻²]	[mm]	[mm]	[mm]	[%]	[years]	m	f	w	nw	
51.2	1.691	-	-	-	17.9	-	14.0	10.8	3.1	29.0	23.3		f	w		DISTANCE-RUNNING
50.2	1.669	0.905	0.936	17.6	18.0	-0.4	24.0	20.2	3.8	18.6	27.4		f	w		RUNNING
58.0	1.790	0.940	1.024	18.3	18.1	0.2	5.8	4.1	1.7	41.5	18.1	m		w		CYCLING
55.2	1.740	-	-	-	18.2	-	29.8	26.6	3.2	11.9	20.8		f	w		MID-DISTANCE-RUNNING
54.9	1.700	-	-	-	19.0	-	55.0	47.4	7.6	16.1	21.7		f	w		DISTANCE-RUNNING
49.2	1.600	0.860	0.920	19.0	19.2	-0.3	28.1	24.8	3.3	13.4	17.3		f	w		GYM ARTISTIC
57.4	1.728	0.958	0.926	18.4	19.2	-0.8	42.8	36.3	6.5	17.9	27.3		f	w		CYCLING
56.4	1.706	0.913	0.942	19.2	19.4	-0.2	43.1	39.5	3.6	9.0	18.8		f		nw	SOCCER
55.9	1.678	0.833	0.982	21.2	19.9	1.3	17.2	15.3	2.0	12.8	21.9		f	w		ATLETISM
65.6	1.817	0.940	0.993	20.4	19.9	0.5	18.8	13.7	5.2	37.8	19.3	m			nw	SWIMMING
67.6	1.830	0.960	1.020	20.4	20.2	0.2	60.2	53.7	6.5	12.0	28.9		f		nw	SWIMMING
57.5	1.684	0.917	0.937	19.7	20.3	-0.5	38.3	35.7	2.6	7.1	18.9		f		nw	SWIMMING
61.5	1.740	-	-	-	20.3	-	30.8	26.6	4.1	15.5	20.3		f		nw	SWIMMING
70.3	1.855	0.977	0.985	20.6	20.4	0.1	46.4	40.1	6.3	15.7	27.0	m			nw	TENIS
69.2	1.835	-	-	-	20.6	-	8.7	6.5	2.2	33.4	22.8	m		w		TRIATHLON
59.4	1.690	0.900	0.990	20.7	20.8	-0.1	18.5	15.5	3.0	19.3	21.8		f		nw	SWIMMING
66.3	1.784	0.965	0.973	20.4	20.8	-0.4	7.6	6.1	1.5	24.5	18.7	m		w		CYCLING
53.8	1.605	0.878	0.864	20.2	20.9	-0.7	44.1	38.9	5.3	13.5	25.3		f	w		SWIMMING
57.9	1.663	0.872	0.929	21.2	20.9	0.2	48.0	42.8	5.2	12.2	31.8	m		w		LONG-DISTANCE-RUNNING
61.3	1.700	0.980	0.950	19.5	21.2	-1.7	26.1	23.4	2.7	11.5	27.7		f		nw	SWIMMING
66.1	1.760	0.940	-	21.2	21.3	-0.2	6.4	5.2	1.2	23.3	19.7	m		w		GYM ARTISTIC
49.6	1.522	0.814	0.855	21.2	21.4	-0.2	27.9	24.3	3.6	15.0	32.2		f	w		RUNNING
66.8	1.764	0.950	0.980	21.1	21.5	-0.3	46.5	44.0	2.6	5.9	19.8		f		nw	SWIMMING
65.5	1.740	0.910	-	21.9	21.6	0.3	16.7	13.7	3.0	21.8	24.0	m		w		GYM ARTISTIC
60.4	1.663	0.899	0.918	21.4	21.8	-0.4	29.1	25.5	3.6	14.0	34.2		f	w		ROAD RACE CYCLING
63.9	1.710	0.900	-	21.9	21.9	0.2	11.7	9.8	1.9	19.2	22.5	m		w		GYM ARTISTIC
62.2	1.687	0.934	0.873	20.9	21.9	-0.9	58.2	55.2	3.0	5.5	27.0		f		nw	SWIMMING
60.3	1.660	0.890	0.950	21.6	21.9	-0.3	33.1	29.2	3.9	13.3	19.3		f	w		GYM ARTISTIC
72.0	1.810	-	-	-	22.0	-	24.8	19.3	5.5	28.4	25.9	m		w		TRIATHLON
72.3	1.809	0.950	0.978	22.3	22.1	0.2	22.6	18.1	4.4	24.4	22.5	m			nw	JUDO
60.3	1.652	0.882	0.945	21.9	22.1	-0.2	44.9	37.8	7.2	18.9	21.9		f	w		PENTATHLON
70.8	1.790	-	-	-	22.1	-	17.2	14.9	2.3	15.6	19.3	m		w		TRIATHLON
61.7	1.670	0.900	0.930	21.8	22.1	-0.4	44.6	39.4	5.2	13.2	22.2		f	w		GYM ARTISTIC
58.2	1.619	0.860	0.867	22.2	22.2	0.0	51.1	43.1	8.0	18.7	19.3		f	w		SWIMMING
55.3	1.570	0.850	0.860	22.0	22.4	-0.5	36.9	32.6	4.4	13.4	22.4		f	w		GYM ARTISTIC
55.3	1.570	0.849	0.860	22.0	22.4	-0.4	56.2	52.1	4.1	7.8	18.7		f		nw	SWIMMING
60.5	1.642	0.870	0.864	22.4	22.4	0.0	6.3	5.4	1.0	18.3	21.5	m		w		ATLETISM
59.2	1.620	0.899	0.905	21.5	22.6	-1.0	87.4	81.7	5.7	7.0	27.8		f			SWIMMING
68.8	1.745	0.898	0.959	23.3	22.6	0.7	40.9	33.2	7.7	23.2	34.3	m		w		TRIATHLON
76.7	1.840	0.940	1.080	23.5	22.7	0.8	20.8	18.7	2.2	11.7	25.5	m			nw	HOCKEY ICE
69.2	1.744	0.919	0.990	22.9	22.8	0.1	14.2	11.1	3.1	27.7	20.4	m			nw	SOCCER
74.8	1.813	-	-	-	22.8	-	18.6	15.6	3.0	19.5	20.8	m			nw	SWIMMING
70.9	1.763	0.940	0.941	22.7	22.8	-0.1	68.7	62.3	6.4	10.3	20.9		f		nw	TENIS
74.1	1.799	0.941	1.034	23.2	22.9	0.3	9.5	7.8	1.8	22.5	22.6	m		w		CROSS COUNTRY SKIING
68.5	1.723	0.928	0.952	22.7	23.1	-0.4	7.9	5.8	2.1	36.0	32.8	m		w		CYCLING
64.5	1.670	0.901	0.931	22.7	23.1	-0.4	8.8	6.8	2.1	30.4	20.1	m			nw	SOCCER
64.8	1.673	0.950	0.909	21.6	23.2	-1.5	70.7	64.7	6.1	9.4	19.5		f		nw	SOCCER
77.2	1.824	0.944	1.045	23.8	23.2	0.6	80.0	72.5	7.5	10.3	18.1		f		nw	VOLLEYBALL
75.1	1.790	0.950	1.020	23.4	23.4	0.0	7.2	5.8	1.4	23.3	26.8	m		w		SNOWBOARDING
80.8	1.856	0.970	1.026	23.8	23.5	0.3	25.8	19.1	6.7	35.0	32.9	m		w		TRIATHLON
71.3	1.742	0.957	0.967	22.7	23.5	-0.8	48.4	44.1	4.3	9.8	23.8	m			nw	SOCCER
64.9	1.660	0.896	0.839	23.1	23.6	-0.4	76.4	73.3	3.1	4.2	22.8		f		nw	FIELD HOCKEY
83.8	1.882	1.021	1.031	23.1	23.7	-0.5	52.3	45.4	6.9	15.1	21.4	m			nw	SOCCER
83.0	1.870	1.000	1.070	23.5	23.7	-0.2	9.1	6.9	2.3	33.3	26.4	m			nw	SWIMMING
71.5	1.733	-	-	-	23.8	-	83.0	77.2	5.8	7.5	19.5		f		nw	WATERPOLO
77.2	1.800	0.970	1.010	23.4	23.8	-0.4	23.6	21.3	2.4	11.0	28.3	m			nw	HOCKEY ICE
85.0	1.885	-	-	-	23.9	-	11.8	9.6	2.2	23.0	25.0	m		w		TRIATHLON
86.8	1.899	1.008	1.063	24.0	24.1	0.0	50.5	42.8	7.7	18.0	23.8	m		w		CYCLING
70.8	1.710	-	-	-	24.2	-	56.8	51.9	4.8	9.3	20.1		f		nw	WATERPOLO
66.0	1.650	0.870	-	24.4	24.2	0.1	8.7	7.0	1.7	24.5	24.0	m		w		GYM ARTISTIC
69.0	1.683	-	-	-	24.4	-	52.2	47.1	5.2	11.0	24.3		f		nw	KAYAK
90.0	1.920	-	-	-	24.4	-	15.4	13.2	2.2	16.6	23.0	m			nw	SWIMMING
71.6	1.710	0.933	0.954	23.8	24.5	-0.7	80.0	75.8	4.1	5.5	19.3		f		nw	VOLLEYBALL
80.0	1.800	0.960	1.040	24.5	24.7	-0.2	34.8	31.1	3.7	11.9	23.7	m			nw	HOCKEY ICE
67.8	1.657	0.884	0.853	24.5	24.7	-0.2	158.8	149.9	8.9	5.9	20.3		f		nw	BASKETTBALL
71.4	1.700	-	-	-	24.7	-	35.5	29.6	5.9	19.7	25.2	m		w		TRIATHLON
75.6	1.738	0.917	0.913	25.1	25.0	0.1	64.6	55.9	8.7	15.6	29.2		f		nw	TRIATHLON
72.2	1.696	0.920	0.898	24.5	25.1	-0.6	9.1	6.9	2.2	32.3	27.0	m		w		CYCLING
73.4	1.701	-	-	-	25.4	-	91.0	85.9	5.2	6.0	23.7		f		nw	WATERPOLO
90.1	1.877	1.006	1.008	25.3	25.6	-0.3	8.7	6.9	1.8	26.7	24.1	m		w		ATLETISM
74.5	1.705	-	-	-	25.6	-	91.1	81.5	9.5	11.7	20.2		f		nw	KAYAK
85.4	1.796	0.967	0.964	26.1	26.5	-0.4	43.5	38.0	5.6	14.6	22.3	m			nw	SOCCER
79.0	1.727	0.920	0.920	26.4	26.5	-0.1	114.6	107.1	7.5	7.0	21.0		f		nw	BASKETTBALL
91.4	1.840	0.970	1.077	27.1	27.0	0.1	118.1	113.0	5.2	4.6	23.1		f		nw	ROWING OPEN
95.0	1.845	0.982	0.957	27.8	27.9	-0.1	54.7	46.3	8.4	18.0	21.3	m			nw	BASKETTBALL
98.1	1.839	1.000	1.030	28.3	29.0	-0.7	35.9	31.5	4.4	13.9	18.5	m			nw	RUGBY
68.5	**1.738**	**0.926**	**0.958**	**22.4**	**22.6**	**-0.2**	**39.9**	**35.5**	**4.3**	**17.0**	**23.4**			**MEAN**		
11.07	**0.09**	**0.05**	**0.06**	**2.24**	**2.28**	**0.51**	**29.87**	**28.37**	**2.17**	**8.71**	**4.10**			**SD**		

Table A2

GROUPS		m [kg]	h [m]	s [m]	l [m]	MI [kgm^{-2}]	BMI [kgm^{-2}]	D_I [mm]	D_E [mm]	D_F [mm]	$D_{F,\%}$ [%]	N
ALL	MEAN	68.5	1.738	0.926	0.958	22.4	22.6	39.9	35.5	4.3	14.1	76
	SD	11.1	0.086	0.046	0.063	2.2	2.3	29.9	28.4	2.2	6.1	
ALL_m	MEAN	74.2	1.789	0.946	0.991	23.1	23.1	22.6	19.0	3.6	18.0	37
	SD	10.0	0.073	0.042	0.052	2.1	2.1	15.9	14.1	2.1	5.4	
ALL_f	MEAN	63.2	1.689	0.906	0.929	21.7	22.1	56.2	51.2	5.0	10.3	39
	SD	9.3	0.067	0.042	0.058	2.2	2.3	30.9	29.7	2.0	4.0	
ALL_w	MEAN	64.5	1.720	0.908	0.946	21.8	21.7	24.8	21.0	3.8	17.2	36
	SD	10.5	0.090	0.048	0.061	1.9	2.0	15.8	13.9	2.2	5.0	
ALL_nw	MEAN	72.2	1.754	0.942	0.967	23.0	23.4	53.5	48.7	4.8	11.3	40
	SD	10.4	0.079	0.037	0.064	2.4	2.2	33.0	31.6	2.1	5.7	
ALL_m_w	MEAN	70.9	1.769	0.931	0.979	22.6	22.6	17.6	14.4	3.2	19.7	21
	SD	8.8	0.076	0.044	0.061	1.8	1.8	14.4	12.4	2.2	4.6	
ALL_m_nw	MEAN	78.6	1.815	0.963	1.002	23.7	23.8	29.2	25.1	4.1	15.9	16
	SD	10.0	0.060	0.033	0.044	2.3	2.4	15.8	14.2	1.9	5.8	
ALL_f_w	MEAN	55.6	1.651	0.877	0.913	20.7	20.4	34.8	30.1	4.7	13.7	15
	SD	4.2	0.059	0.038	0.041	1.6	1.7	12.1	10.4	1.8	3.0	
ALL_f_nw	MEAN	67.9	1.713	0.925	0.939	22.4	23.1	69.6	64.4	5.2	8.2	24
	SD	8.4	0.062	0.033	0.065	2.3	2.1	31.7	30.4	2.1	3.1	
C1_m_w	MEAN	65.4	1.715	0.905	-	22.4	22.3	10.9	8.9	1.9	18.1	4
	SD	1.0	0.048	0.029	-	1.4	1.3	4.4	3.7	0.7	1.5	
C2_m_w	MEAN	75.1	1.780	0.936	0.964	23.7	23.5	20.5	16.1	4.3	20.3	4
	SD	13.0	0.109	0.063	0.073	1.2	1.4	16.2	12.9	3.4	4.4	
C3_m_w	MEAN	69.1	1.793	0.944	1.026	21.6	21.5	7.5	5.9	1.6	22.2	3
	SD	9.6	0.005	0.006	0.007	2.9	2.9	1.9	1.8	0.2	6.2	
C4_m_w	MEAN	73.7	1.804	-	-	-	22.7	19.6	16.0	3.6	19.2	5
	SD	6.4	0.068	-	-	-	1.7	10.8	9.1	1.9	4.5	
C5_m_w	MEAN	70.3	1.753	0.939	0.963	22.6	22.8	24.6	20.9	3.7	19.3	5
	SD	10.6	0.093	0.051	0.062	1.8	1.9	22.5	20.0	2.6	6.4	
C1_m_nw	MEAN	79.2	1.828	0.968	1.050	23.7	23.7	22.1	19.5	2.6	14.0	4
	SD	2.9	0.034	0.025	0.032	0.5	0.8	10.5	10.0	0.7	7.3	
C2_m_nw	MEAN	75.8	1.832	0.962	0.978	22.8	22.6	35.6	29.6	6.1	19.0	4
	SD	13.1	0.022	0.020	0.015	3.5	3.7	17.6	16.1	1.7	6.2	
C3_m_nw	MEAN	84.7	1.791	0.979	0.999	25.5	26.3	42.1	37.8	4.3	10.6	2
	SD	19.0	0.069	0.030	0.045	4.0	3.9	8.9	8.9	0.0	2.3	
C4_m_nw	MEAN	82.4	1.867	-	-	-	23.6	17.0	14.4	2.6	15.3	2
	SD	10.7	0.076	-	-	-	1.2	2.2	1.6	0.6	1.5	
C5_m_nw	MEAN	75.7	1.773	0.952	0.979	23.7	24.0	29.7	25.3	4.4	17.7	4
	SD	10.4	0.089	0.054	0.042	1.6	1.7	21.4	19.2	2.2	5.5	
C1_f_w	MEAN	56.6	1.625	0.875	0.915	21.1	21.4	35.7	31.5	4.2	11.7	4
	SD	5.7	0.048	0.024	0.039	1.4	1.5	6.9	6.1	0.8	0.1	
C2_f_w	MEAN	57.1	1.639	0.863	0.915	21.4	21.3	39.3	33.7	5.6	13.7	4
	SD	2.8	0.033	0.022	0.059	0.9	1.1	15.1	12.5	2.7	2.4	
C3_f_w	MEAN	60.4	1.663	0.899	0.918	21.4	21.8	29.1	25.5	3.6	12.3	1
	SD	-	-	-	-	-	-	-	-	-	-	
C4_f_w	MEAN	53.8	1.710	-	-	-	18.4	32.9	28.3	4.7	15.7	3
	SD	2.2	0.026	-	-	-	0.6	20.7	18.3	2.6	6.1	
C5_f_w	MEAN	52.4	1.640	0.892	0.906	19.1	19.6	31.6	26.9	4.6	14.6	3
	SD	4.3	0.106	0.073	0.044	1.9	1.7	9.9	8.4	1.6	1.4	
C1_f_nw	MEAN	60.9	1.698	0.922	0.955	20.6	21.2	40.2	36.2	4.0	11.1	4
	SD	5.1	0.106	0.060	0.070	1.0	1.0	21.0	19.6	1.7	3.7	
C2_f_nw	MEAN	73.3	1.721	0.915	0.907	24.7	24.8	101.7	93.8	7.9	8.7	4
	SD	5.0	0.045	0.023	0.038	1.5	1.5	44.3	43.8	1.2	3.6	
C3_f_nw	MEAN	67.0	1.709	0.928	0.935	22.3	22.8	70.8	67.1	3.7	5.4	6
	SD	12.4	0.080	0.029	0.085	2.6	2.3	29.5	28.3	1.4	1.1	
C4_f_nw	MEAN	70.1	1.712	-	-	-	23.9	67.5	61.7	5.8	9.2	6
	SD	4.6	0.021	-	-	-	1.9	24.7	23.5	1.9	2.8	
C5_f_nw	MEAN	67.5	1.728	0.935	0.963	22.1	22.6	68.4	63.1	5.3	7.9	4
	SD	9.0	0.066	0.016	0.058	2.2	2.2	17.5	16.4	1.8	1.8	

Table A3: Thickness value differences ABS(δ_I) at the individual sites
Absolute values of differences of the three measurers from their mean values for thickness measurements with fibrous structures included (index: "I"). For the expert group (C1 and C2), number of comparisons at each of the eight sites is: N=(16+16)3=96; for novices (C3-C5): N=(12+16+16)3=132. Abbreviations: UA: upper abdomen, LA: lower abdomen, EO: external oblique, DT: distal triceps, BR: brachioradialis, LT: lateral thigh, FT: front thigh, and MC: medial calf. IQR: inter quartile range, Q3: third quartile (First 75% of values when ordered according to increasing values). Data correspond to Figs. 6a and 6b.

C1, C2	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [mm]	0.15	0.19	0.14	0.11	0.15	0.06	0.10	0.11
IQR [mm]	0.22	0.28	0.24	0.17	0.22	0.08	0.14	0.12
Q3 [mm]	0.29	0.36	0.29	0.23	0.31	0.11	0.17	0.19
C3 - C5	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [mm]	0.24	0.47	0.24	0.22	0.22	0.14	0.18	0.19
IQR [mm]	0.37	0.72	0.35	0.27	0.37	0.28	0.30	0.27
Q3 [mm]	0.47	0.88	0.48	0.36	0.45	0.34	0.37	0.35

Table A4: Thickness value differences ABS(δ_E) at the individual sites
Absolute values of differences of the three measurers from their mean values for thickness measurements with fibrous structures excluded (index: "E"). For number of comparisons and abbreviations see Table A3. Data correspond to Figs. 6c and 6d.

C1, C2	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [mm]	0.19	0.19	0.12	0.12	0.16	0.08	0.15	0.13
IQR [mm]	0.24	0.51	0.22	0.15	0.22	0.10	0.19	0.19
Q3 [mm]	0.31	0.59	0.28	0.20	0.28	0.13	0.25	0.26
C3 - C5	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [mm]	0.23	0.51	0.21	0.20	0.22	0.15	0.31	0.22
IQR [mm]	0.41	0.93	0.40	0.24	0.36	0.28	0.39	0.30
Q3 [mm]	0.50	1.10	0.49	0.34	0.47	0.35	0.52	0.40

Table A5: Relative thickness value differences at the individual sites related to the given subcutaneous adipose tissue thickness (d_I)
ABS($\delta_{I,rel}$)=100·ABS(δ_I)/d_I. The table shows the percentages of measurement differences from the three measurers' mean for thickness measurements with fibrous structures included (index: "I"). For number of comparisons and abbreviations see Table A3. Data correspond to Figs. A1a and b.

C1, C2	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [%]	4.61	2.62	6.10	4.35	5.33	4.15	2.12	3.41
IQR [%]	6.64	3.26	12.16	5.87	6.77	10.22	3.07	5.43
Q3 [%]	8.69	4.38	14.84	8.00	8.69	12.24	3.94	7.57
C3 - C5	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [%]	6.53	5.53	12.65	5.99	6.94	10.39	4.27	5.63
IQR [%]	9.44	6.86	21.58	10.93	10.49	17.35	7.26	14.68
Q3 [%]	12.59	9.53	27.36	13.57	12.56	22.05	8.84	17.00

Table A6: Relative thickness value differences at the individual sites related to the given subcutaneous adipose tissue thickness (d_E)

ABS($\delta_{E,rel}$)=100·ABS(δ_E)/d_E. The table shows the percentages of measurement differences from the three measurers' mean for thickness measurements with fibrous structures included (index: "I"). For number of comparisons and for abbreviations see Table A3. Data correspond to Figs. A1c and A1b.

C1, C2	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [%]	4.80	4.21	7.04	4.45	5.39	6.20	3.91	5.94
IQR [%]	9.51	5.93	12.37	4.96	8.59	9.23	4.60	8.92
Q3 [%]	11.96	7.88	15.16	7.20	10.20	12.05	6.17	11.87
C3 - C5	UA	LA	EO	ES	DT	BR	FT	MC
MEDIAN [%]	8.42	9.13	16.02	6.71	7.76	12.39	6.88	8.54
IQR [%]	11.64	9.60	21.98	8.31	14.60	16.81	8.00	15.79
Q3 [%]	15.13	13.87	28.56	11.47	17.84	22.68	11.50	18.94

Fig. A1: Percentages of measurement differences δ_{rel} with respect to the subcutaneous adipose tissue thicknesses (d) at the individual sites.

According data are presented in Tables A5 and A6. Abbreviations. UA: upper abdomen, LA: lower abdomen, EO: external oblique, DT: distal triceps, BR: brachioradialis, LT: lateral thigh, FT: front thigh, and MC: medial calf.
Index "I": fibrous structures (fasciae) included, index "E": fibrous structures excluded.
A1a: Experience examiners (C1,C2): ABS($\delta_{I,rel}$)=100·ABS(δ_I)/d_I.
Number of comparisons at each of the eight sites is: N=(16+16)3=96.
A1b: Novices (C3-5): ABS($\delta_{I,rel}$)=100·ABS(δ_I)/d_I.
Number of comparisons at each of the eight sites is: N=(12+16+16)3=132.
A1c: Experience examiners: ABS (δ_E) for each of the eight sites.
ABS($\delta_{E,rel}$)=100·ABS(δ_E)/d_E. N=96.
A1d: Novices: ABS (δ_E) for each of the eight sites.
ABS($\delta_{E,rel}$)=100·ABS(δ_E)/d_E. N=132.
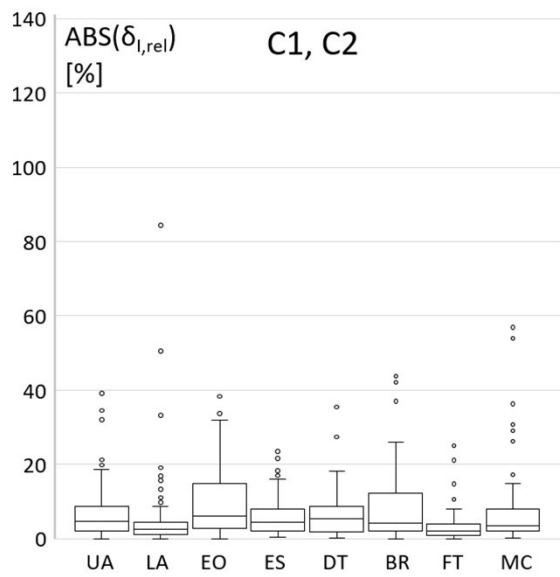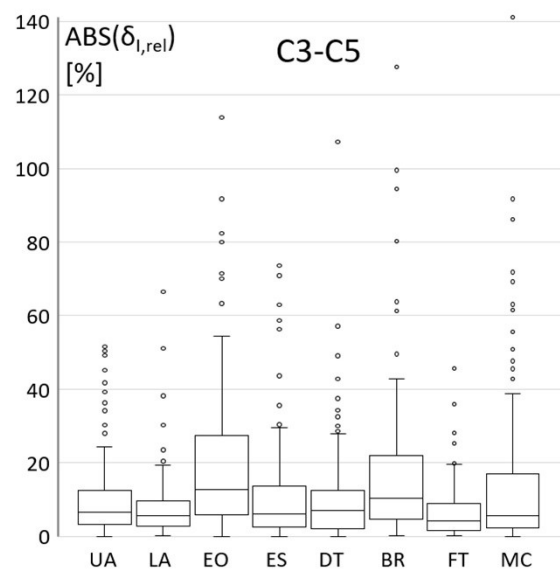
A1a:

A1b:

A1c:
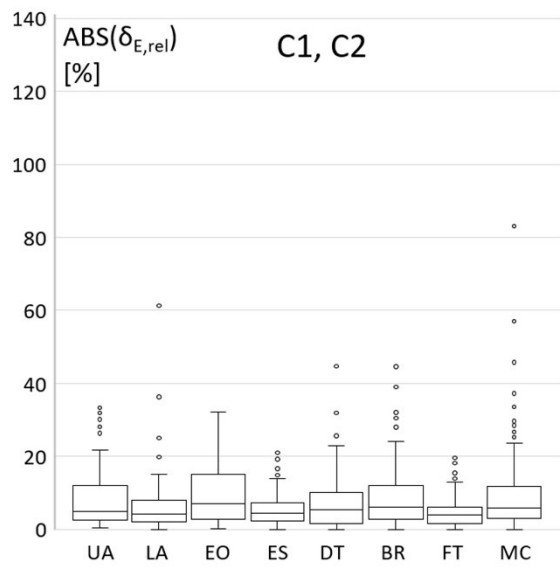
A1d:
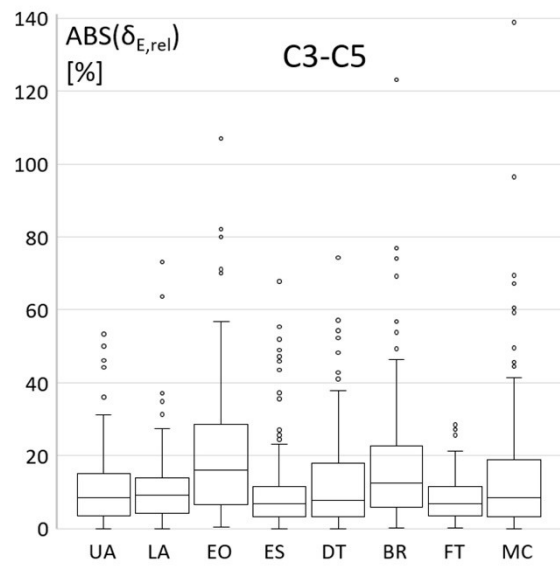
Table A7: Preliminary normative data for the subcutaneous adipose tissue (SAT) sums of the eight standardised sites [12].

Sum of SAT (mm)	Valuation	Comment
Competitive Athletes - Female		
below 25	Extremely low	Medical surveillance recommended
25 to 35	Very low	Surveillance recommended
35 to 50	Low	Desirable range
50 to 70		Noticeable ballast weight
above 70		Considerable ballast weight
Competitive Athletes - Male		
below 12	Extremely low	Medical surveillance recommended
12 to 20	Very low	Surveillance recommended
20 to 30	Low	Desirable range
30 to 50		Noticeable ballast weight
above 50		Considerable ballast weight
General Public - Female		
below 25	Extremely low	Medical surveillance recommended
25 to 35	Very low	Surveillance recommended
35 to 80	Low	Desirable range
80 to 110		Noticeable ballast weight
110 to 140		Considerable ballast weight
140 to 180	High	Surveillance recommended
above 180	Very high	Medical surveillance recommended
General Public - Male		
below 12	Extremely low	Medical surveillance recommended
12 to 20	Very low	Surveillance recommended
20 to 60	Low	Desirable range
60 to 100		Noticeable ballast weight
100 to 130		Considerable ballast weight
130 to 180	High	Surveillance recommended
above 180	Very high	Medical surveillance recommended

Table A8: D_I and D_E in mm of all 76 athletes. M1, M2, M3: measurers of each centre

N	CENTRE	SUBJECT	$D_{I,M1}$	$D_{I,M2}$	$D_{I,M3}$	$D_{I,mean}$	$D_{E,M1}$	$D_{E,M2}$	$D_{E,M3}$	$D_{E,mean}$
1	C1	S1	37.40	36.83	36.55	36.93	32.20	33.16	32.37	32.58
2	C1	S2	44.39	44.70	44.68	44.59	39.07	39.04	40.06	39.39
3	C1	S3	27.65	28.66	28.08	28.13	24.96	25.56	23.92	24.81
4	C1	S4	31.96	32.46	34.79	33.07	28.28	28.66	30.67	29.20
5	C1	S5	6.40	6.05	6.59	6.35	4.78	4.94	5.74	5.15
6	C1	S6	16.35	16.44	17.17	16.65	13.64	13.64	13.72	13.67
7	C1	S7	8.45	8.96	8.65	8.69	7.10	6.75	7.08	6.98
8	C1	S8	11.81	12.01	11.36	11.73	10.11	9.87	9.54	9.84
9	C1	S9	55.86	56.69	55.93	56.16	51.72	53.30	51.25	52.09
10	C1	S10	35.08	34.73	34.51	34.77	30.81	30.21	32.23	31.08
11	C1	S11	20.93	20.55	21.04	20.84	19.02	18.34	18.63	18.66
12	C1	S12	23.15	23.86	23.87	23.63	21.80	21.42	20.63	21.28
13	C1	S13	8.96	8.93	9.51	9.13	6.80	6.60	7.16	6.85
14	C1	S14	18.16	17.87	19.36	18.46	15.10	15.24	16.09	15.48
15	C1	S15	61.91	58.04	60.61	60.19	55.09	52.05	54.09	53.74
16	C1	S16	25.74	26.83	25.71	26.09	22.98	23.66	23.55	23.40
17	C2	S1	26.32	26.05	25.13	25.83	20.35	19.59	17.48	19.14
18	C2	S2	40.60	41.22	40.96	40.93	33.91	32.75	32.97	33.21
19	C2	S3	8.57	8.30	9.31	8.73	6.87	6.62	7.17	6.89
20	C2	S4	44.58	43.50	44.28	44.12	39.91	38.84	37.85	38.87
21	C2	S5	45.73	44.55	44.55	44.94	38.83	38.35	36.20	37.79
22	C2	S6	16.83	17.39	17.49	17.24	15.52	15.41	14.90	15.28
23	C2	S7	6.28	6.32	6.41	6.34	5.56	5.25	5.25	5.35
24	C2	S8	51.28	51.15	50.88	51.10	42.82	43.46	42.90	43.06
25	C2	S9	18.44	19.72	18.35	18.84	13.43	14.89	12.67	13.66
26	C2	S10	22.51	22.42	22.80	22.58	17.56	18.71	18.17	18.15
27	C2	S11	54.94	55.00	54.02	54.65	46.82	46.21	45.87	46.30
28	C2	S12	112.41	116.50	107.55	112.15	101.68	110.46	102.20	104.78
29	C2	S13	63.77	64.66	65.31	64.58	57.64	57.88	52.04	55.85
30	C2	S14	44.78	46.56	47.88	46.41	38.22	40.38	41.71	40.10
31	C2	S15	159.35	160.77	156.15	158.76	149.70	153.28	146.63	149.87
32	C2	S16	67.95	70.18	68.06	68.73	60.31	64.22	62.47	62.33
33	C3	S1	9.14	11.29	8.18	9.54	6.94	9.46	6.98	7.79
34	C3	S2	6.15	7.65	7.76	7.19	5.39	5.88	6.22	5.83
35	C3	S3	30.14	28.58	28.57	29.10	25.97	26.49	24.14	25.53
36	C3	S4	5.75	5.94	5.70	5.80	3.58	4.43	4.29	4.10
37	C3	S5	113.71	125.21	115.42	118.11	106.60	120.84	111.43	112.96
38	C3	S6	59.72	55.74	59.19	58.22	58.16	50.27	57.18	55.20
39	C3	S7	87.06	84.18	90.85	87.36	81.59	79.00	84.43	81.67
40	C3	S8	38.09	36.88	39.90	38.29	34.77	34.45	37.99	35.74
41	C3	S9	80.03	69.54	79.62	76.40	76.68	66.06	77.29	73.34
42	C3	S10	47.30	45.63	46.70	46.54	44.42	43.18	44.25	43.95
43	C3	S11	46.94	49.86	48.39	48.40	43.39	45.11	43.77	44.09
44	C3	S12	35.43	38.71	33.45	35.86	29.49	34.59	30.37	31.48
45	C4	S1	17.94	18.49	15.26	17.23	14.58	17.35	12.77	14.90
46	C4	S2	13.38	12.21	9.69	11.76	10.99	10.55	7.13	9.56
47	C4	S3	25.78	23.20	25.48	24.82	18.00	17.69	22.29	19.33
48	C4	S4	8.97	8.18	8.98	8.71	6.98	6.12	6.48	6.53
49	C4	S5	29.81	31.02	28.58	29.80	26.75	28.84	24.28	26.62
50	C4	S6	58.59	54.44	51.93	54.99	49.38	49.04	43.63	47.35
51	C4	S7	32.97	36.65	36.86	35.49	25.97	31.59	31.37	29.64
52	C4	S8	13.60	14.67	13.64	13.97	10.44	11.85	10.19	10.83
53	C4	S9	28.33	32.42	31.51	30.75	24.45	28.56	26.87	26.63
54	C4	S10	15.20	15.40	15.63	15.41	12.28	14.19	13.20	13.22
55	C4	S11	19.53	19.37	16.86	18.59	16.24	16.13	14.29	15.55
56	C4	S12	97.95	82.32	92.89	91.05	89.65	72.66	82.26	81.52
57	C4	S13	55.85	51.55	49.33	52.24	52.79	45.11	43.26	47.05
58	C4	S14	55.69	54.98	59.58	56.75	47.95	50.37	57.46	51.93
59	C4	S15	82.47	83.22	83.20	82.96	75.45	76.40	79.73	77.19
60	C4	S16	90.95	89.35	92.71	91.00	86.50	82.95	88.09	85.85
61	C5	S1	49.74	45.31	48.96	48.00	44.92	39.85	43.59	42.79
62	C5	S2	43.19	39.72	45.39	42.77	36.31	33.68	38.89	36.29
63	C5	S3	30.61	25.22	27.97	27.93	27.49	21.83	23.56	24.29
64	C5	S4	25.57	22.04	24.36	23.99	22.63	17.34	20.71	20.23
65	C5	S5	8.30	7.33	7.24	7.62	6.62	5.97	5.78	6.12
66	C5	S6	53.85	46.70	51.02	50.52	47.88	37.69	42.93	42.83
67	C5	S7	7.98	8.71	7.00	7.90	5.67	6.22	5.53	5.81
68	C5	S8	8.22	9.91	9.15	9.09	6.98	6.20	7.44	6.87
69	C5	S9	80.62	76.31	83.00	79.98	77.19	71.79	78.55	75.84
70	C5	S10	71.87	66.91	73.37	70.72	65.19	60.26	68.51	64.65
71	C5	S11	45.05	40.25	45.26	43.52	39.73	34.51	39.65	37.96
72	C5	S12	43.68	41.04	44.43	43.05	41.38	36.74	40.31	39.48
73	C5	S13	52.05	50.47	54.37	52.30	47.00	42.22	47.05	45.42
74	C5	S14	81.56	78.33	80.11	80.00	73.58	70.64	73.33	72.52
75	C5	S15	10.26	7.43	8.72	8.80	7.11	6.04	7.09	6.75
76	C5	S16	15.05	12.67	14.80	14.17	12.77	9.63	10.90	11.10