

# Precision of handheld multispectral optoacoustic tomography for muscle imaging

Alexandra L. Wagner<sup>a</sup>, Vera Danko<sup>a</sup>, Anna Federle<sup>b</sup>, Daniel Klett<sup>b</sup>, David Simon<sup>c</sup>, Rafael Heiss<sup>d</sup>, Jörg Jüngert<sup>a</sup>, Michael Uder<sup>d</sup>, Georg Schett<sup>c</sup>, Markus F. Neurath<sup>b</sup>, Joachim Woelfle<sup>a</sup>, Maximilian J. Waldner<sup>b</sup>, Regina Trollmann<sup>a</sup>, Adrian P. Regensburger<sup>a,1</sup>, Ferdinand Knieling<sup>a,\*,1</sup>

<sup>a</sup> Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany

<sup>b</sup> Department of Medicine 1, University Hospital Erlangen, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany

<sup>c</sup> Department of Medicine 3, University Hospital Erlangen, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany

<sup>d</sup> Institute of Radiology, University Hospital Erlangen, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany

## ARTICLE INFO

### Keywords:

MSOT  
Multispectral optoacoustic tomography  
Muscle imaging  
Optoacoustics  
Photoacoustics

## ABSTRACT

Photo-or optoacoustic imaging (OAI) allows quantitative imaging of target tissues. Using multi-wavelength illumination with subsequent ultrasound detection, it may visualize a variety of different chromophores at centimeter depth. Despite its non-invasive, label-free advantages, the precision of repeated measurements for clinical applications is still elusive. We present a multilayer analysis of  $n = 1920$  imaging datasets obtained from a prospective clinical trial (NCT03979157) in  $n = 10$  healthy adult volunteers. All datasets were analyzed for 13 single wavelengths (SWL) between 660 nm–1210 nm and five MSOT-parameters (deoxygenated/oxygenated/total hemoglobin, collagen and lipid) by a semi-automated batch mode software. Intraclass correlation coefficients (ICC) were good to excellent for intrarater (SWL: 0.82–0.92; MSOT-parameter: 0.72–0.92) and interrater reproducibility (SWL: 0.79–0.87; MSOT-parameter: 0.78–0.86), with the exception for MSOT-parameter lipid (interrater ICC: 0.56). Results were stable over time, but exercise-related effects as well as inter- and intramuscular variability were observed. The findings of this study provide a framework for further clinical OAI implementation.

## 1. Introduction

Within recent years, photoacoustic or optoacoustic imaging (PAI/OAI) has been evolving from a preclinical to a promising future clinical imaging modality [1,2]. OAI allows non-invasive, quantitative imaging of target tissues with high spatial resolution by means of near-infrared laser-induced thermoelastic expansion and subsequent detection of ultrasonic waves [3].

As comparable to clinical ultrasound (US), handheld OAI devices can offer structural, functional and kinetic information, either generated by endogenous absorbers or exogenous contrast agents [4]. The use of a multi-wavelength illumination, termed as multispectral optoacoustic tomography (MSOT), may allow to resolve endogenous chromophores

such as deoxygenated and oxygenated hemoglobin, melanin and lipids [3,5,6]. The potential clinical impact of label-free imaging has been shown in various studies, including the visualization of disease activity in Crohn's disease [7], malignant features in human breast tissue [8], microvascular dysfunction in systemic sclerosis [9], diverse cardiovascular applications [10] and muscle degeneration in Duchenne muscular dystrophy [11]. From a clinical perspective, MSOT offers non-invasive functional imaging without the need for ionizing radiation or long scanning times.

Since MSOT enables quantification, standardization and measures of precision are needed to evaluate the accuracy and reliability of the technical systems and to ensure quality and control of measured parameters for further studies and subsequent clinical translation [12]. In

\* Corresponding author at: Pediatric Experimental and Translational Imaging Laboratory (PETI-Lab), Department of Pediatrics and Adolescent Medicine, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Loschgestraße 15, 91054, Erlangen, Germany.

E-mail address: [ferdinand.knieling@uk-erlangen.de](mailto:ferdinand.knieling@uk-erlangen.de) (F. Knieling).

<sup>1</sup> These authors contributed equally

preclinical studies, OAI was reported to consistently perform equal or better (with variations of smaller than 10 %) than other preclinical imaging modalities, underlining its clinical potential [13]. Efforts are currently made, as seen with the foundation of the International Photoacoustics Standardization Consortium (IPASC), to reach an international consensus on PAI standardization [14].

To address the need of further clinical imaging precision, we provide a large-scale semi-automated multidimensional analysis of OAI imaging data derived from a prospective single-center clinical study in healthy adult volunteers. We demonstrate the precision of quantitative imaging measurements for intra- and interrater reproducibility, temporal and exercise-associated errors as well as inter- and intramuscular variability.

## 2. Material and methods

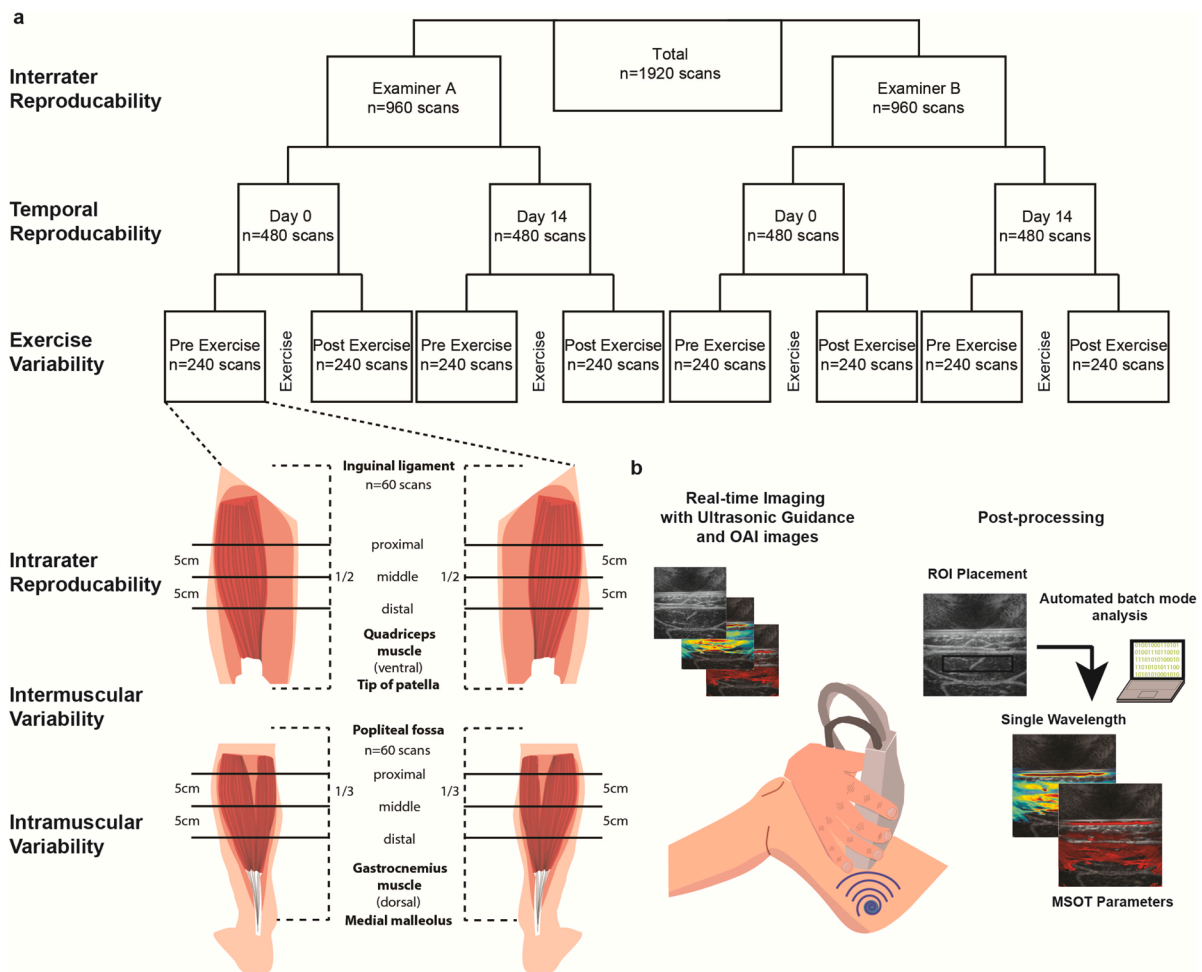
### 2.1. Study design and participants

A prospective, monocentric, single arm clinical trial (investigator-

initiated trial) was performed. The ethics committee of the University of Erlangen-Nürnberg granted approval for the study (number: 169\_19B) and it was registered at clinicaltrials.gov (Identifier: NCT03979157). All volunteers gave written informed consent and the study was conducted in accordance with the declaration of Helsinki. Imaging before and after exercise was performed in the same sequence at two visits in a 14-day-interval.  $n = 10$  healthy volunteers ( $n = 5$  women,  $n = 5$  men)  $> 18$  years of age were recruited. All participants were asked to shave the imaging sites prior to examination. Exclusion criteria included anamnestic or clinical signs of myopathy, pregnancy and tattoos on the skin area to be examined.

### 2.2. Study flow

After inclusion, a resting period after arrival at the study site was ensured. MSOT imaging was performed according to the following protocol: all imaging procedures were consecutively repeated by two different, independent examiners (investigator A, two years of clinical



**Fig. 1.** Study flow and MSOT principle.

a) Schematic study flow. All scans (total  $n = 1920$ ) were performed by two independent examiners (interrater reproducibility) in a 14-day interval (temporal reproducibility) before and after muscular activity through exercise (6 MWT) (exercise variability). Each scan was repeated twice (intrater reproducibility). Quadriceps muscles and gastrocnemius muscles (intermuscular variability) were examined on both sides at three different positions: proximal, middle and distal (intramuscular variability). The middle position was determined as follows: quadriceps muscles: sitting position, 1/2 between inguinal ligament and tip of patella; gastrocnemius muscles: standing position, 1/3 between popliteal fossa and middle malleolus. Proximal and distal scans were performed 5 cm above or below the middle scan, respectively. 6MWT = 6 minute walk test.

b) The concept of MSOT. A transducer is placed on the skin and pulsed multi-wavelength laser light illuminates underlying tissue. Based on the photoacoustic effect, absorption of light leads to thermoelastic expansion and subsequent ultrasonic emissions. Due to specific absorption and scattering of incident light, different chromophores (hemoglobin, lipid, and collagen) can be spectrally unmixed. MSOT signals for SWLs (660 nm, 680 nm, 715 nm, 730 nm, 760 nm, 800 nm, 850 nm, 920 nm, 1000 nm, 1030 nm, 1064 nm, 1100 nm, and 1210 nm) and MSOT parameters (deoxygenated hemoglobin (Hb), oxygenated hemoglobin (HbO<sub>2</sub>), total hemoglobin (HbT), collagen, and lipid) were obtained. MSOT = Multispectral Optoacoustic Tomography, SWL = Single wavelength, ROI = region of interest.

MSOT imaging experience and B, no previous experience, both with experience in standard US). MSOT was performed on two anatomical muscle groups of the leg (upper: quadriceps, lower: gastrocnemius) on both sides (left and right). MSOT scans were taken at three different positions (proximal, middle and distal) with two repetitive scans per position. Identical scanning positions were ensured by skin markings according to anatomical landmarks at each visit (middle quadriceps: half between inguinal ligament and tip of patella; proximal: +5 cm; distal: -5 cm; middle gastrocnemius: one third of popliteal fossa to middle malleolus; proximal: +5 cm; distal: -5 cm). All MSOT scans were acquired in an identical fashion before and after standardized exercise (6-minute walk test (6MWT) [15,16]). The entire procedure was repeated after 14 days (Fig. 1a).

### 2.3. MSOT technical details

A prototype hybrid reflected-ultrasound computed tomography (RUCT)/MSOT imaging system (MSOT Acuity Echo, iThera Medical GmbH, Munich, Germany) comprising a handheld 2D detector (4 MHz, 256 transducer elements) was used. The detector had a field of view of 40 mm x 40 mm and a spatial resolution of 150  $\mu$ m. The probe was positioned at about 90 degree angle on the skin, coupled by transparent US gel (medimex GmbH, Germany). RUCT was used for guidance during image acquisition. For laser safety all patients and examiners wore appropriate laser safety goggles.

### 2.4. Data analysis

MSOT signals were obtained at 660 nm, 680 nm, 715 nm, 730 nm, 760 nm, 800 nm, 850 nm, 920 nm, 1000 nm, 1030 nm, 1064 nm, 1100 nm, 1210 nm and reconstructed using backprojection algorithm [17]. A polygonal region of interest (ROI) was traced just beneath the muscle fascia according to the RUCT image by an independent blinded reader using cLabs software (V2.65, iThera Medical GmbH). Quantification was performed by iLabs software (iThera Medical GmbH) in a batch-processing mode (Fig. 1b). All MSOT signals are given in arbitrary units (a.u.) and correspond to the mean signal value within the ROI. For evaluation of MSOT unmixing parameters multiple SWL from the near- (MSOT hemoglobin) and extended near (MSOT collagen and MSOT lipid) infrared spectrum of light were used for further analyses to depict specific spectral signatures from multispectral measurements ([5,7,11, 18]). Deoxygenated hemoglobin (Hb), oxygenated hemoglobin (HbO<sub>2</sub>), collagen and lipid were spectrally unmixed using linear regression algorithm [3,19]. 715 nm, 730 nm, 760 nm, 800 nm and 850 nm were used for unmixing Hb and HbO<sub>2</sub>, whereas collagen and lipid were unmixed using 680 nm, 715 nm, 730 nm, 760 nm, 800 nm, 850 nm, 920 nm, 1000 nm, 1030 nm, 1064 nm and 1100 nm. Total hemoglobin (HbT) signal was calculated as  $HbT = Hb + HbO_2$ . In the following, all described parameters implicit MSOT derived parameters, e.g. HbT = MSOT parameter HbT or collagen = MSOT parameter collagen.

### 2.5. Statistical analysis

Given the pilot study design, no case number calculation could be performed. The data are mean  $\pm$  SD and tested for normal distribution using the Shapiro-Wilk test. Analyses were performed in a paired fashion. For two normal distributed groups, a paired *t*-test, and for nonparametric data, Wilcoxon matched-pair signed rank tests were applied. Analysis of variance (ANOVA) for three or more groups and in case of nonparametric distribution Friedman Test was used. Correlations were calculated using Pearson (*r*) or nonparametric Spearman correlation coefficient (*r<sub>s</sub>*) as appropriate. To define the upper and lower limit, the 95 %CI was used. The intraclass correlation coefficient (ICC) was performed [20], by a two-way mixed effects, absolute agreement, single measurement for intrarater reproducibility and two-way random effects, absolute agreement, multiple measurements for interrater

reproducibility [21]. Grading was as follows: < 0.4 were considered as poor, 0.41–0.60 moderate, 0.61–0.80 good, 0.81–1 very good and 1 as perfect agreement [22]. Bland Altman plots were calculated including 95 % limits of agreement [23]. A *p* < 0.05 was considered as statistically significant. Statistical analysis was carried out using IBM SPSS Statistics (Version 24, International Business Machines Corporation, USA) and GraphPad Prism (Version 8, GraphPad Software, Inc., La Jolla, CA, USA).

## 3. Results

### 3.1. Study population

*n* = 10 healthy volunteers (*n* = 5 women, *n* = 5 men) were included in the clinical trial between October 14th and November 11th, 2019. Age was  $25.7 \pm 4.4$  years, weight was  $67.0 \pm 15.1$  kg, height was  $174.0 \pm 10.2$  cm, body-mass-index (BMI) was  $21.9 \pm 2.8$  kg/m<sup>2</sup>. No serious adverse events were reported (Table 1).

### 3.2. Intrarater reproducibility

All scans (*n* = 1920) were used to assess the reproducibility within one examiner. Intraclass correlations (ICC) between scan 1&2 were excellent for all SWL (range: 0.82 – 0.92, Fig. 2a, Supplementary Table 1). The derived MSOT parameters showed good/excellent ICCs for Hb (0.91; 95 %CI: 0.90–0.92), HbO<sub>2</sub> (0.87; 95 %CI: 0.85–0.88), HbT (0.92; 95 %CI: 0.91–0.93), collagen (0.82; 95 %CI: 0.79–0.84) and lipid (0.72; 95 %CI: 0.68–0.76) (Fig. 2b, Supplementary Table 2). The experienced (examiner A, range: 0.83–0.93) showed a minimally better intrarater reproducibility than the non-experienced examiner (examiner B, range: 0.81–0.91). Bland-Altman plots did not show a large dispersion (bias for 800 nm: -0.43, HbT:  $-1.12 \times 10^{-4}$  and collagen: -0.08, respectively, Fig. 2c).

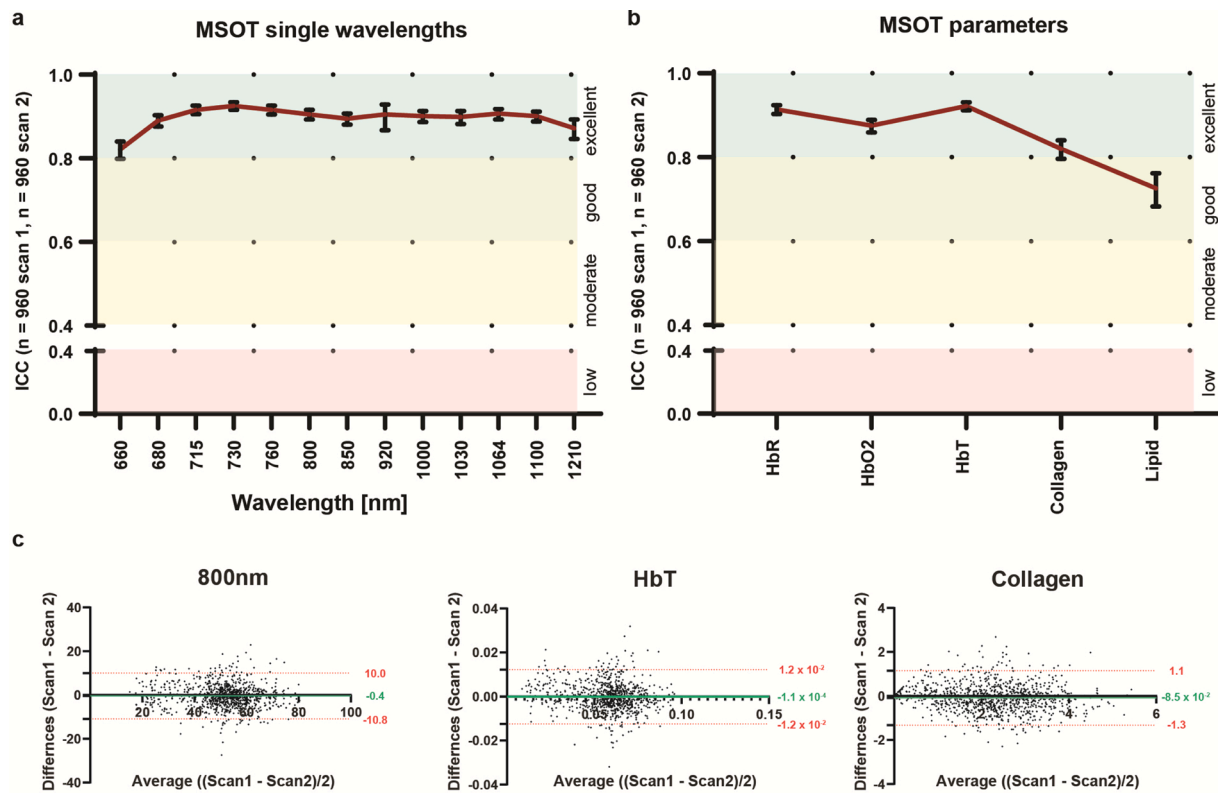
### 3.3. Interrater reproducibility

The following analyses were performed with the mean values of scan 1 (*n* = 960) and 2 (*n* = 960). Values were matched for each anatomic position. ICCs were excellent for SWLs 660 nm–1100 nm (range: 0.82 – 0.87, for 95 %CI) and good for 1210 nm (0.79; 95 %CI: 0.75–0.82) (Fig. 3a, Supplementary Table 3). For MSOT parameters, ICCs were

**Table 1**  
Healthy volunteers' clinical characteristics.

Healthy volunteers clinical characteristics		
	Female ( <i>n</i> = 5)	Male ( <i>n</i> = 5)
Age, years	25.0 $\pm$ 2.7	26.4 $\pm$ 5.8
Height, cm	165.4 $\pm$ 5.3	182.6 $\pm$ 4.7
Weight, kg	55.6 $\pm$ 8.0	78.4 $\pm$ 11.2
BMI, kg/m <sup>2</sup>	20.2 $\pm$ 1.8	23.5 $\pm$ 2.9
Medication	Contraceptive ( <i>n</i> = 2)	none
	Sertraline ( <i>n</i> = 1)	
MSOT scanning time/visit 1*, min	49.0 $\pm$ 5.5	52.0 $\pm$ 9.1
MSOT scanning time/visit 2*, min	43.0 $\pm$ 2.7	48.0 $\pm$ 2.7
Walking distance/visit 1**, m	655.0 $\pm$ 97.5	615.0 $\pm$ 91.2
Walking distance/visit 2**, m	685.0 $\pm$ 51.8	635.0 $\pm$ 112.6
Adverse events	none	Uncomfortable feeling of coolness of ultrasound gel ( <i>n</i> = 1)

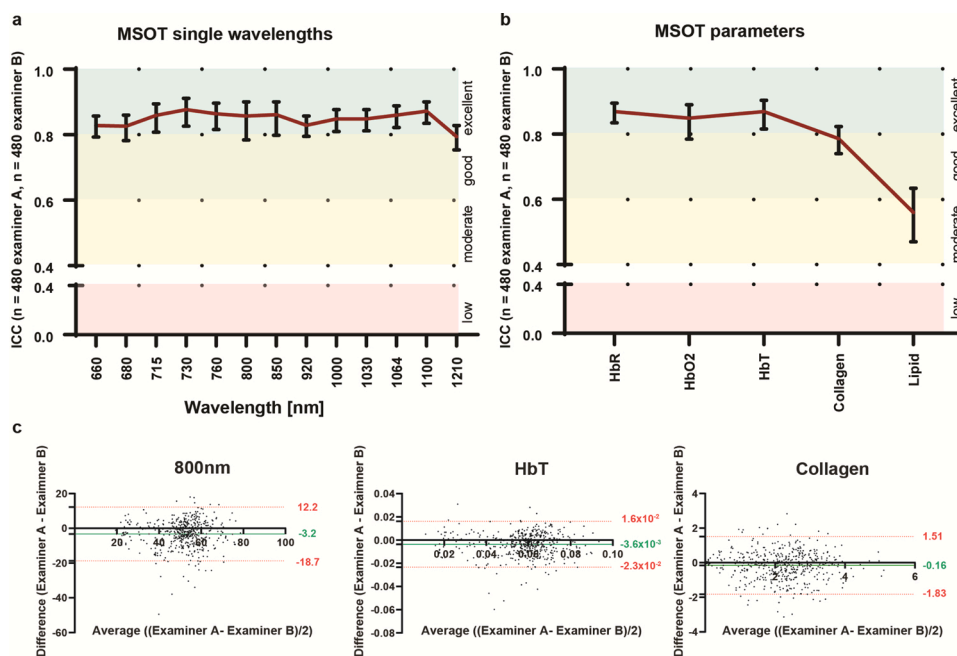
Categorical variables are provided as numbers. Continuous variables are mean  $\pm$  SD. \*approximated to 5 min, \*\*approximated to 25 m. MSOT scanning time/visit comprises the time for the whole assessment including both measurements before and after muscular activity and the 6 MWT itself. 6MWT = 6 minute walk test.



**Fig. 2.** Intratester reproducibility.

a + b) Intraclass correlation coefficients (ICCs, two-way mixed effects, absolute agreement, single measurement) were calculated between scan 1 and 2 (n = 960 scan 1, n = 960 scan 2, n = 120 independent muscle regions, n = 10 independent subjects) for 13 single wavelengths (a) and MSOT parameters (b) deoxygenated hemoglobin (Hb), oxygenated hemoglobin (HbO2), total hemoglobin (HbT), collagen and lipid. The grading of Landis and Koch was applied: ICC values < 0.4 were considered as poor, 0.41–0.60 moderate, 0.61–0.80 good, 0.81–1 very good and 1 as perfect agreement [22]. For detailed information see Supplementary Table 1 + 2.

c) Bland-Altman analysis illustrating the difference between the paired measurement (scan 1 – scan 2) and the average of the measurements ((scan 1 – scan 2)/2) (n = 960 scan 1, n = 960 scan 2, n = 120 independent muscle positions, n = 10 independent subjects) for 800 nm, HbT and collagen. The green line indicates the bias, the red lines lower and upper 95 % limits of agreement.



**Fig. 3.** Interrater reproducibility.

a + b) Intraclass correlation coefficients (ICCs, two-way random effects, absolute agreement, multiple measurements) were calculated between examiner A and B (n = 480 examiner A, n = 480 examiner B, n = 120 independent muscle positions, n = 10 independent subjects) for 13 single wavelengths (a) and MSOT parameters (b) deoxygenated hemoglobin (Hb), oxygenated hemoglobin (HbO2), total hemoglobin (HbT), collagen and lipid. The grading of Landis and Koch was applied: ICC values < 0.4 were considered as poor, 0.41–0.60 moderate, 0.61–0.80 good, 0.81–1 very good and 1 as perfect agreement [22]. For detailed information see Supplementary Table 3 + 4.

c) Bland-Altman analysis illustrating the difference between the paired measurement (examiner A – examiner B) and the average of the measurements ((examiner A - examiner B)/2) (n = 480 examiner A, n = 480 examiner B, n = 120 independent muscle positions, n = 10 independent subjects) for 800 nm, HbT and collagen. The green line indicates the bias, the red lines lower and upper 95 % limits of agreement.



excellent for Hb (0.86; 95 %CI: 0.83–0.89), HbO<sub>2</sub> (0.84; 95 %CI: 0.78–0.89) and HbT (0.86; 95 %CI: 0.81–0.90), good for collagen (0.78; 95 %CI: 0.74–0.82) and moderate for lipid (0.56; 95 %CI: 0.47–0.63) (Fig. 3b, Supplementary Table 4). Bland-Altman plots did not show a large dispersion (bias for 800 nm:  $-3.26$ , HbT:  $-3.60 \times 10^{-3}$  and collagen:  $-0.15$ ) (Fig. 3c).

As intra- and interrater results for each SWL and MSOT parameters were in similar ranges, the following results are presented for 800 nm, HbT and collagen.

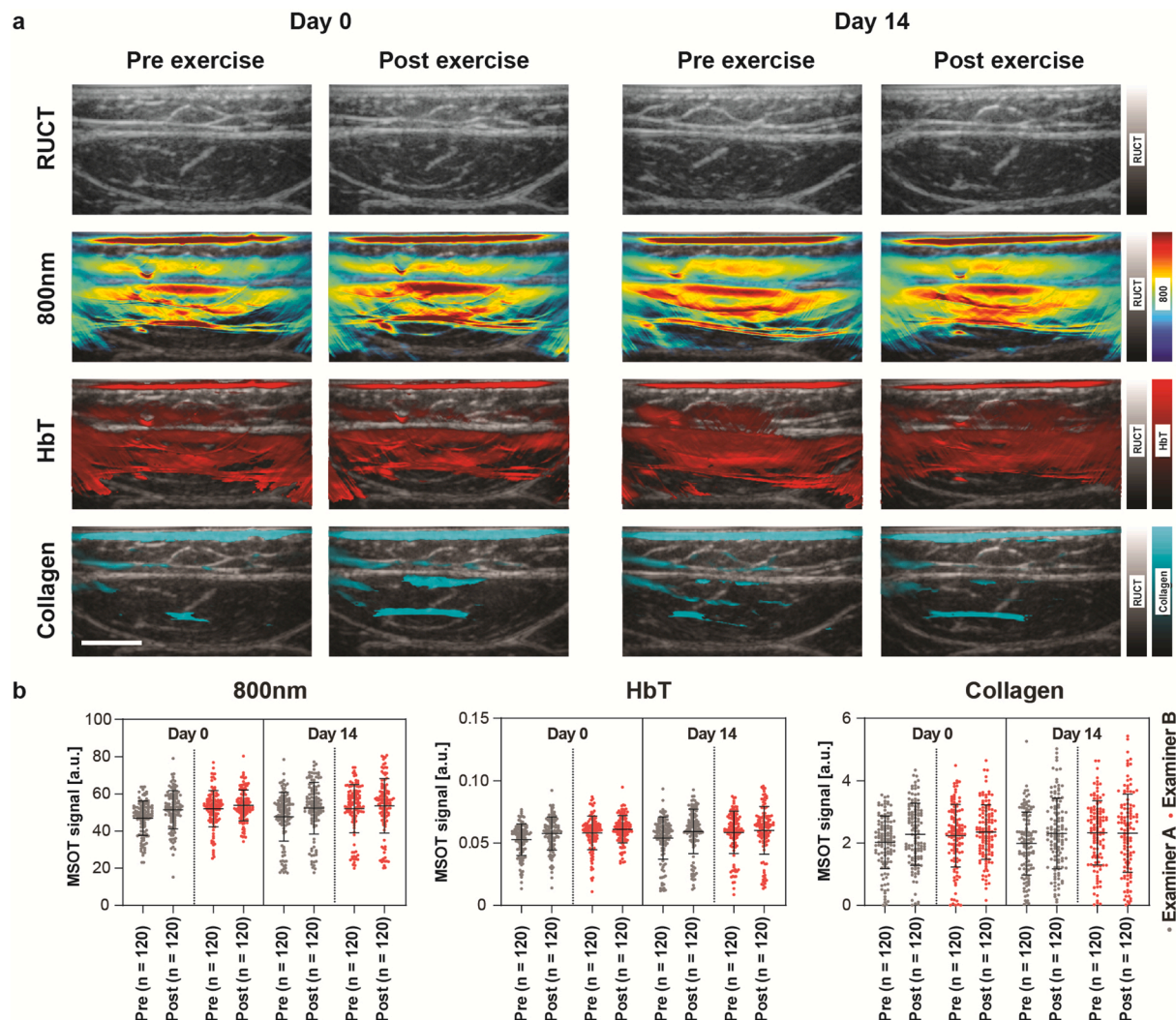
### 3.4. Temporal reproducibility

The scans of two visits (day 0 and 14) were compared (Fig. 4a). Mean values of scan 1 and 2 were used and pre-exercise ( $n = 960$ , mean  $n = 480$ ) were analyzed separately from post-exercise scans ( $n = 960$ , mean  $n = 480$ ) to limit confounding of variables by muscular activity (Fig. 4b). Results were similar comparing day 0 and 14 without any significant difference for SWL 800 nm (pre-exercise:  $p = 0.3639$ ; post exercise:  $p = 0.6487$ ), HbT (pre-exercise:  $p = 0.1918$ ; post exercise:  $p = 0.7009$ ) and collagen (pre-exercise:  $p = 0.8947$ ; post exercise:  $p = 0.8647$ ). Similar

results were retrieved when analyzing each examiner individually (Supplementary Table 5a).

### 3.5. Muscular activity

The means of scan 1 and 2 of all pre- ( $n = 960$  total scans, mean  $n = 480$ ) and all post-exercise scans (total  $n = 960$ , mean  $n = 480$ ) were compared (Fig. 4a + b). Differences were significant for 800 nm (pre vs. post:  $49.72 \pm 11.63$  a.u. vs.  $52.86 \pm 12.09$  a.u.,  $p < 1.0 \times 10^{-15}$ ), HbT ( $5.59 \times 10^{-2} \pm 1.53 \times 10^{-2}$  a.u. vs.  $5.96 \times 10^{-2} \pm 1.57 \times 10^{-2}$  a.u.,  $p < 1.0 \times 10^{-15}$ ) and collagen ( $2.14 \pm 0.98$  a.u. vs.  $2.31 \pm 1.07$  a.u.,  $p = 2.84 \times 10^{-5}$ ). Analyzing each examiner separately, differences were consistently significant for examiner A. For examiner B, 10–20 min post-exercise, findings were still significant for 800 nm (pre vs. post:  $52.07 \pm 11.46$  a.u. vs.  $53.77 \pm 11.91$  a.u.,  $p = 0.0006$ ), HbT ( $5.84 \times 10^{-2} \pm 1.53 \times 10^{-2}$  a.u. vs.  $6.06 \times 10^{-2} \pm 1.56 \times 10^{-2}$  a.u.,  $p = 0.0005$ ) but not for collagen ( $2.28 \pm 1.01$  a.u. vs.  $2.33 \pm 1.08$  a.u.,  $p = 0.2558$ ) (Supplementary Table 5b).



**Fig. 4.** Temporal reproducibility and muscular activity.

a) Exemplary images of the right quadriceps muscle of one patient (female, BMI 17.4 kg/m<sup>2</sup>) on day 0 and day 14 before and after exercise, respectively. Displayed are RUCT images (RUCT) for muscular guidance and superimposed MSOT signals (800 nm, HbT, Collagen). Scale bar indicates 1 cm. RUCT = reflection ultrasound computed tomography, HbT = total hemoglobin.

b) Displayed are  $n = 120$  independent muscle positions from  $n = 10$  independent subjects at each timepoint from examiner A (grey) and examiner B (red) for 800 nm, HbT and collagen. Each dot represents the mean (scan 1 and scan 2) MSOT signal per independent muscle position and subject. Data are given as mean value with standard deviation. Pre = pre-exercise, post = post-exercise, HbT = total hemoglobin.

### 3.6. Intermuscular differences

Generally higher MSOT signal levels were found for gastrocnemius muscles for 800 nm (quadriceps vs. gastrocnemius muscles:  $46.10 \pm 11.43$  a.u. vs.  $56.48 \pm 10.09$  a.u.,  $p < 1.0 \times 10^{-15}$ ), HbT ( $5.51 \times 10^{-2} \pm 1.54 \times 10^{-2}$  a.u. vs.  $6.40 \times 10^{-2} \pm 1.30 \times 10^{-2}$  a.u.,  $p < 1.0 \times 10^{-15}$ ) and collagen ( $1.85 \pm 0.89$  a.u. vs.  $2.60 \pm 1.03$  a.u.,  $p < 1.0 \times 10^{-15}$ ) (Fig. 5a). This effect was found to be more pronounced in women (HbT:  $4.27 \times 10^{-2} \pm 1.65 \times 10^{-2}$  a.u. vs.  $6.15 \times 10^{-2} \pm 1.49 \times 10^{-2}$  a.u.,  $p < 1.0 \times 10^{-15}$ ) than in men (HbT:  $6.02 \times 10^{-2} \pm 0.70 \times 10^{-2}$  a.u. vs.  $6.65 \times 10^{-2} \pm 1.02 \times 10^{-2}$  a.u.,  $p < 5.9 \times 10^{-14}$ ) (Fig. 5b). A negative correlation of MSOT HbT signals with muscle depth was found ( $r_s = -0.6795$ , % CI  $-0.7130$  to  $-0.6428$ ,  $p < 1.0 \times 10^{-15}$ ) (Fig. 5c, Supplementary Fig. 1). In accordance with this finding of depth limitation, variability of MSOT HbT signal increased with muscle depth (Supplementary Fig. 2).

### 3.7. Intramuscular variability

To discriminate potential changes along the muscles, the proximal, middle and distal parts were examined (each scan twice;  $n = 6$  scans/muscle,  $n = 10$  independent subjects,  $n = 960$  means from  $n = 1920$  total scans). Significant intramuscular differences for 800 nm, HbT and collagen, except for middle vs. distal scans for 800 nm ( $p = 0.1502$ ) were found. When comparing all mean values of proximal, middle and distal positions, overall maximal difference was 7.5 % for 800 nm and 5.4 % for HbT and 22.7 % for collagen (Fig. 6a).

### 3.8. Gender differences

MSOT signals were statistically significantly lower in women for quadriceps (HbT women:  $4.27 \times 10^{-2} \pm 0.15 \times 10^{-3}$  a.u. vs. men:

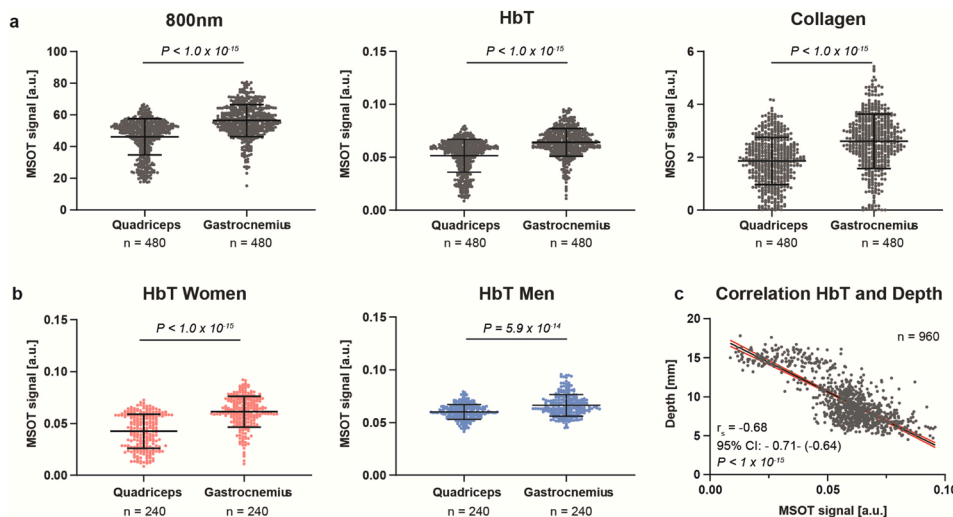
$6.02 \times 10^{-2} \pm 0.17 \times 10^{-2}$  a.u.,  $p = 0.0034$ ) and for gastrocnemius muscles ( $6.15 \times 10^{-2} \pm 0.21 \times 10^{-2}$  a.u. vs.  $6.65 \times 10^{-2} \pm 0.36 \times 10^{-2}$  a.u.,  $p = 0.0371$ ). The mean depth for quadriceps was  $12.90 \pm 2.59$  mm for women and  $8.25 \pm 1.23$  mm for men, for gastrocnemius muscles depth was  $9.66 \pm 2.72$  mm vs.  $7.19 \pm 1.14$  mm. Intramuscular differences were evident in gastrocnemius, but not uniformly in quadriceps muscles (Fig. 6b). Overall differences of all mean values of proximal, middle and distal scans separated for muscle region (quadriceps and gastrocnemius muscle) and gender were in the range of 0.6–10.1 % for HbT (Fig. 6c, Supplementary Table 6).

## 4. Discussion

This study showed the reproducibility of a large quantity of clinical MSOT imaging data sets for various confounding variables, using a semi-automated batch mode software as analyzing tool. Next to intra- and interrater reproducibility this is, to the best of our knowledge, the first clinical MSOT study investigating temporal and exercise variability as well as inter- and intramuscular differences.

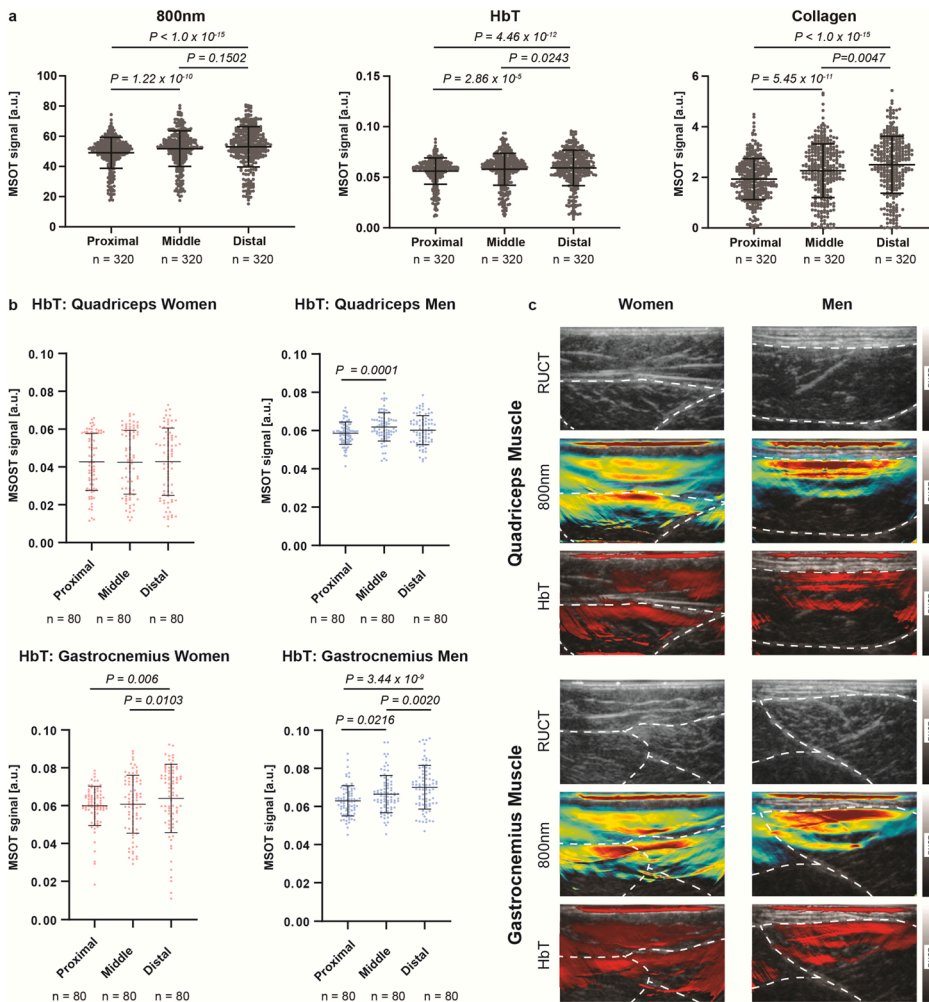
We demonstrate high intra- and interrater ICCs for all SWLs and spectrally unmixed parameters except lipid, with good feasibility also for a less experienced examiner. This correlation is in the range of or even better than comparable handheld ultrasound systems [24–26], especially for muscle imaging applications [27–29]. The lower ICC values for lipid might be explainable by generally low lipid levels in healthy muscles, especially in comparable lean individuals [30], so that little variations in probe positioning may have high impact on signal quantifications.

As hypothesized, muscular activity via exercise (6MWT) and suggested consecutive increased blood flow in muscles led to higher MSOT signals. Interestingly, this seems to be a transient effect, as after 10–20



**Fig. 5.** Intermuscular variability.

a) Displayed are the MSOT signals of quadriceps muscles and gastrocnemius muscles ( $n = 480$  quadriceps muscles/  $n = 480$  gastrocnemius muscles, from  $n = 120$  independent muscle positions, from  $n = 10$  independent subjects) for 800 nm, HbT and collagen. Each dot represents the mean MSOT signal (scan 1 and scan 2) of one independent muscle position in an individual subject. For analyses a paired, non-parametric  $t$ -test (Wilcoxon matched-pair signed rank test) was used. Data are given as mean value with standard deviation.  $P \leq 0.05$  was considered as statistically significant. HbT = total hemoglobin. b) Analogous to 4a), the same analyses were performed separately for women (red) and men (blue) ( $n = 240$  quadriceps muscles/  $n = 240$  gastrocnemius muscles for women and men each, from  $n = 120$  independent muscle positions, from  $n = 10$  independent subjects) shown for HbT. For analysis a paired, non-parametric  $t$ -test (Wilcoxon matched-pair signed rank test) was used. Data are given as mean value with standard deviation.  $P \leq 0.05$  was considered as statistically significant. HbT = total hemoglobin. c) Correlation between MSOT signal, and the muscle depth is shown ( $n = 960$  data points, each dot representing the mean MSOT signal (scan 1 and scan 2) of one independent muscle position (quadriceps muscles/ gastrocnemius muscles),  $n = 120$  independent muscle regions;  $n = 10$  independent subjects). Analyses were performed with the non-parametric Spearman correlation ( $r_s$ ). Confidence interval (95 %) is shown in red.  $P \leq 0.05$  was considered as statistically significant. HbT = total hemoglobin.

**Fig. 6.** Intramuscular variability.

a) Displayed are  $n = 320$  dots per proximal, middle and distal muscle position for  $n = 120$  independent muscle positions from  $n = 10$  independent subjects for 800 nm, HbT and collagen. Analyses were performed with the non-parametric Friedman test for multiple comparisons. Data are given as mean value with standard deviation.  $P \leq 0.05$  was considered as statistically significant. HbT = total hemoglobin.

b) Subdivision of each muscle position proximal, middle and distal of quadriceps muscles and gastrocnemius muscles for women (red) and men (blue) ( $n = 80$  for each column,  $n = 120$  independent muscle positions;  $n = 10$  independent subjects). Analyses were performed with ANOVA and the non-parametric Friedman test, both for multiple comparisons. Data are given as mean value with standard deviation.  $P \leq 0.05$  was considered as statistically significant. HbT = total hemoglobin.

c) Representative images of right middle quadriceps muscle and right middle gastrocnemius muscle of a woman (BMI 22.2 kg/m<sup>2</sup>, left) and man (BMI 21.6 kg/m<sup>2</sup>, right), illustrating the difference in subcutaneous fat content and light absorption. Displayed are RUCT images and superimposed MSOT signals for 800 nm and HbT. Muscle fascia are plotted with dashed white lines. Scale bar indicates 1 cm. HbT = total hemoglobin.

min post exercise, differences become smaller and approach baseline pre-exercise values. These transient exercise-induced changes are in agreement with those detected by other techniques like MRI [31,32]. Notably, the effect of muscular activity and the range of signal levels were stable over time (14-day interval), indicating the potential utility for clinical monitoring studies.

When comparing intermuscular imaging signals between two anatomical regions, we found generally higher signals in the gastrocnemius muscle. As penetration depth is a well-known limitation of optical imaging [33], the greater thickness of subcutaneous fat tissue over the thigh might be accountable for this finding [34]. Mean intramuscular signal values between the proximal, middle and distal position of the quadriceps and gastrocnemius muscles were also mostly different, while overall errors between mean values remained considerably low. Intramuscular differences might be explained due to the biological and interindividual variability, personal physical activity levels, sexual differences (e.g. muscle volume) and non-uniform architecture throughout the volume and along the muscles of the lower leg [35–39].

Further analyses revealed relevant gender differences with ubiquitous lower MSOT signal levels in women. Gender variability and heterogeneity of muscles is well described with regard to muscle fiber-type [37,40], fiber bundle length, enzyme activities [37], muscle thickness [35] and volume [38]. It is known that women present with about 10 % more body fat compared to men with the same BMI [41,42]. As our results indicated a relevant difference in thickness of subcutaneous fat between both genders and a negative correlation between MSOT signal and muscle depth, we suggest that limits in penetration depth mainly

account for these gender-specific findings.

To implement MSOT imaging in the clinical routine, challenges as depth limitation and therefore difficulties when investigating deeper anatomical regions need to be addressed. The present technology needs continuative technical development and adaption of the underlying algorithms. Further study-specific limitations include the lack of imaging data of the upper extremity, the short time interval, the effect of more strenuous exercise than the 6MWT on MSOT signals and the small number of subjects due to the pilot feasibility character of the study. In addition, we did not include pathologic muscular changes, where more pronounced differences between healthy and diseased subjects were previously reported [11]. Apart from this, additional studies will be needed to address similar questions in different target tissues.

This study emphasizes the stability and reproducibility of MSOT imaging based on a clinically relevant and representative comprehensive data set. For a wide range of clinical applications, but especially for degenerative or inflammatory muscle diseases in children and adults, our results will serve as a basis for the future establishment of appropriate imaging protocols.

#### Author contributions

ALW, APR and FK conceived the idea and design of the study. RT and FK were the principal investigators. ALW, APR and FK recruited participants. ALW and APR imaged with assistance of VD, AF and DK. ROIs were traced by VD and reviewed by FK. Data collection was completed by ALW. ALW, APR and FK analyzed the data sets. ALW, RH, JJ, MU, GS,



MFN, JW, MJW, APR and FK interpreted the data. ALW wrote the first draft of the manuscript. The manuscript was critically reviewed by all authors.

### Declaration of Competing Interest

APR, MJW and FK are co-inventors with iThera Medical GmbH on an EU patent application (no. EP 19 163 304.9). APR, MJW and FK received travel support by iThera Medical GmbH, Germany. RH, MU and FK report lecture fees from Siemens Healthcare GmbH and APR from Sanofi Genzyme outside the submitted work. Given his role as Section Editor of this journal, FK had no involvement in the peer-review of articles for which he was an author and had no access to information regarding their peer-review. Full responsibility for the peer-review process for this article was delegated to another Editor. All other authors declare no competing interests.

### Acknowledgements

Supported by the Interdisciplinary Center for Clinical Research (FK, APR, DS) and funded by ELAN Fond (P055, APR) at the University Hospital of FAU Erlangen-Nürnberg. Funded by the Else Kröner-Fresenius-Stiftung (Else Kröner-Memorial-Stipendium, no. 2018\_EKMS.03, FK; no. 2019\_EKMS.27, DS). Funded by Frieda Marohn Foundation (RT) and Graduate School in Advanced Optical Technologies (MJW) and the Emerging Fields Initiative of FAU (MFN, DS, GS). MJW and MFN funded by the German Research Foundation (FOR2438 and TRR241). Funded by EU H2020 research and innovation program (no. 830965, MFN, MJW, FK). DS and GS funded by German Research Foundation (FOR2886 & CRC1181), the Federal Ministry for Education and Research (project METARTHROS&MASCARA), H2020 GA 810316 - 4D-Nanoscope ERC Synergy Project, IMI funded project RTCure. The present work was performed in fulfillment of the requirements for obtaining the degree "Dr. med." for A.L.W. at the Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and in (partial) fulfillment of the requirements for obtaining the degree "Dr. rer. biol. hum." for F.K. at the Friedrich-Alexander-University Erlangen-Nürnberg (FAU). We thank Koray Tascilar for statistical consultation.

### Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.pacs.2020.100220>.

### References

- [1] I. Steinberg, et al., Photoacoustic clinical imaging, *Photoacoustics* 14 (2019) 77–98.
- [2] A.P. Regensburger, et al., Shedding light on pediatric diseases: multispectral optoacoustic tomography at the doorway to clinical applications, *Mol. Cell. Pediatr.* 7 (1) (2020) 3.
- [3] V. Ntziachristos, D. Razansky, Molecular imaging by means of multispectral optoacoustic tomography (MSOT), *Chem. Rev.* 110 (5) (2010) 2783–2794.
- [4] A.B.E. Attia, et al., A review of clinical photoacoustic imaging: current and future trends, *Photoacoustics* 16 (2019) 100144.
- [5] S. Tzoumas, et al., Unmixing molecular agents from absorbing tissue in multispectral optoacoustic tomography, *IEEE Trans. Med. Imaging* 33 (1) (2014) 48–60.
- [6] A. Taruttis, V. Ntziachristos, Advances in real-time multispectral optoacoustic imaging and its applications, *Nat. Photonics* 9 (4) (2015) 219–227.
- [7] F. Knieling, et al., Multispectral optoacoustic tomography for assessment of Crohn's disease activity, *N. Engl. J. Med.* 376 (13) (2017) 1292–1294.
- [8] G. Diot, et al., Multispectral optoacoustic tomography (MSOT) of human breast Cancer, *Clin. Cancer Res.* (2017).
- [9] M. Masthoff, et al., Multispectral optoacoustic tomography of systemic sclerosis, *J. Biophotonics* 11 (11) (2018) e201800155.
- [10] A. Karlas, et al., Cardiovascular optoacoustics: from mice to men - a review, *Photoacoustics* 14 (2019) 19–30.
- [11] A.P. Regensburger, et al., Detection of collagens by multispectral optoacoustic tomography as an imaging biomarker for Duchenne muscular dystrophy, *Nat. Med.* 25 (12) (2019) 1905–1915.
- [12] A. Helfen, et al., Multispectral optoacoustic tomography: intra- and interobserver variability using a clinical hybrid approach, *J. Clin. Med.* 8 (1) (2019).
- [13] J. Joseph, et al., Evaluation of precision in optoacoustic tomography for preclinical imaging in living subjects, *J. Nucl. Med.* 58 (5) (2017) 807–814.
- [14] S. Bohndiek, Addressing photoacoustics standards, *Nat. Photonics* 13 (5) (2019) 298.
- [15] A.T.S.Co.P.Sf.C.P.F. Laboratories, ATS statement: guidelines for the six-minute walk test, *Am. J. Respir. Crit. Care Med.* 166 (1) (2002) 111–117.
- [16] C.M. McDonald, et al., The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy, *Muscle Nerve* 41 (4) (2010) 500–510.
- [17] M. Xu, L.V. Wang, Universal back-projection algorithm for photoacoustic computed tomography, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71 (1 Pt 2) (2005) 016706.
- [18] Q. Cao, et al., Multispectral imaging in the extended near-infrared window based on endogenous chromophores, *J. Biomed. Opt.* 18 (10) (2013) 101318.
- [19] D. Razansky, et al., Multispectral opto-acoustic tomography of deep-seated fluorescent proteins in vivo, *Nat. Photonics* 3 (2009) 412–417.
- [20] K.O. McGraw, S.P. Wong, Forming Inferences About Some Intraclass Correlation Coefficients Psychological Methods, 1, 1996, pp. 30–46 (1).
- [21] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, *J. Chiropr. Med.* 15 (2) (2016) 155–163.
- [22] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, *Biometrics* 33 (1) (1977) 159–174.
- [23] D. Giavarina, Understanding Bland altman analysis, *Biochem. Med. (Zagreb)* 25 (2) (2015) 141–151.
- [24] E. Pardo, et al., Reliability of ultrasound measurements of quadriceps muscle thickness in critically ill patients, *BMC Anesthesiol.* 18 (1) (2018) 205.
- [25] K.W.P. Ng, et al., Reliability of bedside ultrasound of limb and diaphragm muscle thickness in critically ill children, *Muscle Nerve* 59 (1) (2019) 88–94.
- [26] F. Bobadilla, et al., Pre-surgical high resolution ultrasound of facial basal cell carcinoma: correlation with histology, *Cancer Imaging* 8 (2008) 163–172.
- [27] W. Nijholt, et al., The reliability and validity of ultrasound to quantify muscles in older adults: a systematic review, *J. Cachexia Sarcopenia Muscle* 8 (5) (2017) 702–712.
- [28] M. Taghipour, et al., Reliability of real-time ultrasound imaging for the assessment of trunk stabilizer muscles: a systematic review of the literature, *J. Ultrasound Med.* 38 (1) (2019) 15–26.
- [29] Mechelli Filippo, Arendt-Nielsen Lars, Stokes Maria, Agyapong-Badu Sandra, Inter-rater and intra-rater reliability of ultrasound imaging for measuring quadriceps muscle and non-contractile tissue thickness of the anterior thigh, *Biomed. Phys. Eng. Express* 5 (2019), 037002.
- [30] P. Malenfant, et al., Fat content in individual muscle fibers of lean and obese subjects, *Int. J. Obes. Relat. Metab. Disord.* 25 (9) (2001) 1316–1321.
- [31] H.M. Nagaraj, et al., Determining exercise-induced blood flow reserve in lower extremities using phase contrast MRI, *J. Magn. Reson. Imaging* 27 (5) (2008) 1096–1102.
- [32] F. Adelnia, et al., Diffusion-weighted MRI with intravoxel incoherent motion modeling for assessment of muscle perfusion in the thigh during post-exercise hyperemia in younger and older adults, *NMR Biomed.* 32 (5) (2019) e4072.
- [33] Oraevsky, A., Jacques, S., Esenaliev, R., Tittel, F., Laser-based optoacoustic imaging in biological tissues., in *SPIE, OE/LASE' 94*, 1994, Los Angeles, CA, United States. 1994.
- [34] S.A. Al-Attar, et al., Semi-automated segmentation and quantification of adipose tissue in calf and thigh by MRI: a preliminary study in patients with monogenic metabolic syndrome, *BMC Med. Imaging* 6 (2006) 11.
- [35] R.S. Chow, et al., Sonographic studies of human soleus and gastrocnemius muscle architecture: gender variability, *Eur. J. Appl. Physiol.* 82 (3) (2000) 236–244.
- [36] A.M. Agur, et al., Documentation and three-dimensional modelling of human soleus muscle architecture, *Clin. Anat.* 16 (4) (2003) 285–293.
- [37] J.A. Simoneau, C. Bouchard, Human variation in skeletal muscle fiber-type proportion and enzyme activities, *Am. J. Physiol.* 257 (4 Pt 1) (1989) E567–72.
- [38] A. Foure, et al., Diffusion properties and 3D architecture of human lower leg muscles assessed with ultra-high-Field-Strength diffusion-tensor MR imaging and tractography: reproducibility and sensitivity to sex difference and intramuscular variability, *Radiology* 287 (2) (2018) 592–607.
- [39] R.L. Lieber, J. Friden, Functional and clinical significance of skeletal muscle architecture, *Muscle Nerve* 23 (11) (2000) 1647–1666.
- [40] E. Nygaard, Skeletal muscle fibre characteristics in young women, *Acta Physiol. Scand.* 112 (3) (1981) 299–304.
- [41] A.S. Jackson, et al., The effect of sex, age and race on estimating percentage body fat from body mass index: the heritage family study, *Int. J. Obes. Relat. Metab. Disord.* 26 (6) (2002) 789–796.
- [42] K. Karastergiou, et al., Sex differences in human adipose tissues - the biology of pear shape, *Biol. Sex Differ.* 3 (1) (2012) 13.





Dr. Wagner is a resident and clinical scientist at the Department of Pediatric and Adolescent Medicine at the University Hospital Erlangen. Her current clinical and research focus is on neuro-pediatrics with special interest in neuromuscular diseases. She works on the implementation of novel non-invasive technologies for the diagnosis of diseases and the monitoring of treatments.



Dr. Knieling is a clinician scientist and group leader at the Department of Pediatric and Adolescent Medicine at the University Hospital Erlangen. His research focuses on the discovery of biological insights by novel imaging modalities and their translation to clinical applications.



Dr. Regensburger is a resident and clinician scientist in the Department of Pediatric and Adolescent Medicine at the University Hospital Erlangen. His current clinical and research focus is on new imaging modalities in medicine with a special focus to light and sound based technologies.