J Musculoskelet Neuronal Interact 2023; 23(1):4-25

Journal of Musculoskeletal and Neuronal Interactions JMNI

Original Article

Mechanographic analysis of the timed 4 stair climb test – methodology and reference data of healthy children and adolescents

David C. Schorling[1], Rainer Rawer[2], Imke Kuhlmann[1], Cornelia Müller[1], Astrid Pechmann[1], Janbernd Kirschner[1]

[1]Department of Neuropediatrics and Muscle Disorders, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Germany;
[2]Novotec GmbH, Pforzheim, Germany

Abstract

Objectives: The timed 4 stair climb test (4SC) is an accepted and widely used tool to assess motor function of patients with neuromuscular diseases. We aimed to establish reference data for the 4SC, and for mechanographic analysis of ascent (4SC-Up) and descent (4SC-Dn) in healthy children and adolescents. **Methods**: We used a custom-made staircase measuring device to assess force, power and velocity during the ascent of 4 stairs in healthy subjects. Secondary outcome measures included mechanographic analyses such as the Chair-Rising-test and the myometric Grip Force-test. **Results**: Data of 288 participants aged 4 to 16 years (144 males, 144 females) were analyzed. A simple algorithm integrating the minimal applied force was used to compensate for different movement strategies. Percentiles for average power, force and horizontal velocity were calculated. While results of the 4SC-Up test showed no age or gender dependency, we found 4SC-Dn results to be age dependent. Mean device measured times were significantly shorter than manually measured times (mean difference -0.19 s; $p<0.001$). **Conclusions**: Mechanographic analysis of the 4SC appears to be a promising tool for evaluation of muscle strength and function of the lower extremities as it enables physically exact measurements of a highly relevant activity of daily living.

Keywords: Biomarkers, Gait Analysis, Stair Climbing, Reference Values

Introduction

Different assessments are available to measure muscular strength and/or muscular function of patients with neuromuscular diseases. These can be divided into instrument-based and non-instrument-based methods. Non-instrument-based methods include manual testing of muscle strength in different groups of muscles following standardized scales (e.g. Medical Research Council scale, MRC) or the use of disease-specific scoring-systems in which accomplishment in various motor tasks are calculated to an overall score reflecting the functional level [e.g. Hammersmith Functional Motor Scale Extended for patients with spinal muscular atrophy (HFSME)[1] or the North Star Assessment for Ambulatory patients with Duchenne muscular dystrophy (NSAA)[2]]. The 6-minute walk has become widely used as a further non-apparatus-based measuring method, in which the maximum walking distance in meters is measured in a

David Christof Schorling participated in workshops sponsored by Biogen and Roche, received compensations for consulting services from Pfizer. Rainer Rawer is employed at the company Novotec GmbH, which manufactures and distributes the devices used in this study. Astrid Pechmann received compensation for presentations and training activities from Biogen, received research funding from Biogen and Novartis. Janbernd Kirschner received research funding and/or compensation for presentations and consulting services from Biogen, Novartis, Pfizer, PTC, Roche and Sarepta.The remaining authors have nothing to declare.

Corresponding author: David C. Schorling MD, Department of Neuropediatrics and Muscle Disorders, Medical Center – University of Freiburg, Germany. Mathildenstrasse 1, 79106 Freiburg, Germany
E-mail: David.Schorling@uniklinik-freiburg.de

Edited by: G. Lyritis
Accepted 12 December 2022

time of 6 minutes under standardized conditions, allowing evaluation of muscle function and endurance. Though shown to be reliable methods, these tools either depend much on the experience of the examiner[3,4] or require some time and effort of patients[5]. Shorter timed function tests (TFT) include the measurement of the time needed to rise from the floor (rise time), the time needed to stand up from sitting position, or the time to climb 4 stairs (4 stair climb, 4SC).

Timed stair tests by manual time measurement have a long-standing relevance in the assessment of patients with neuromuscular and neurological diseases, as well as in geriatric patients as they reflect strength and function of the proximal muscles of the lower extremities, knee extensors and plantar flexors[6–9]. Stair-bound function tests are easily feasible and were shown to have an excellent test-retest reliability in both, adult and paediatric cohorts[10,11]. The 4SC is a reduced form of longer stair tests that is manageable for patients with muscle diseases. In different neuromuscular diseases, the initial symptoms are a noticeable gait pattern and difficulties in climbing stairs, which are often reflecting a weakness of the proximal muscles of the lower extremities. This is characteristic, among others, of Duchenne muscular dystrophy (DMD), where the 4-stair climb time is used in both, clinical routine[12] but also in clinical trials[13–17]. Regarding pediatric cohorts, reference data are available only for defined age groups, especially for younger boys aged 2 to 5 years[18].

On the other hand, apparatus-based measuring methods may be used for assessment of muscle strength and function. The use of myometric tools allows physically exact measurement of strength of single muscle groups. However, a statement about the resulting muscle function is rather difficult as muscle strength and function reflect different qualities - e.g. an individual may compensate the loss of strength in a distinct muscle by compensatory involvement of other muscle groups to achieve sufficient function[19]. With the relatively new method of mechanography, movements are analysed based on the resulting ground reaction forces, using devices with built-in force sensors such as force plates. These systems typically allow the spatially resolved recording of dynamic ground reaction forces. Thereby, analysis of muscle function during the performance of everyday movements such as walking, jumping or getting up from a sitting based on physical parameters like force, power and velocity is possible[20–22,24]. In addition, there are also measuring devices consisting of staircases with built-in force sensors, which allow the analysis of stair ascent and descent in adults and children and determination of the physical quantities force, power and speed. However, the step height of these devices does not correspond with widely used analogous staircase devices with a standardized step height of 15 cm (6').

Given the high relevance of this motion sequence for daily life activities[23] we aimed to establish reference data not only for the timed 4SC, but also for mechanographic analysis during ascent and descent of four steps in children and adolescents aged 4 to 16 years. We further correlated these results with yet available mechanographic and myometric analysis. For this purpose, we used a custom build staircase device with built-in force sensors in each corner corresponding with dimensions in widely used analogous devices.

Materials and Methods

We performed a monocentric, prospective diagnostic study investigating mechanographic measurements in healthy children and adolescents. The study was registered at the German registry for clinical trials (DRKS; DRKS00015240) and approved by the regional ethics committee.

Outcome measures: Primary outcome measure was mechanographic analysis of force, power and velocity during the 4-step Stair Climb-test (4SC-Up=ascent and 4SC-Dn=descent) using a custom-made staircase measuring device, based on the commercially available Leonardo Stair C Mechanograph© (Novotec Medical GmbH, Pforzheim, Germany)[29,30]. The model differs from the original version regarding the height and depth of steps and thus matches with widely used analogous staircase devices with a step height of 15 cm (6') and with requirements for subitems of standardized physiotherapeutic assessments such as the North Star Assessment for ambulant patients with DMD (items 6 to 9). Technical data of the device are displayed in Table S1 and an image of the device as well as further information are placed in the appendix (Figure S1).

Secondary outcome measures included analysis of maximum power during the Chair-Rising-test (CRT) using a Leonardo Mechanograph® GRFP LT device and a bench adjustable for individual body height and analysis of the maximum Grip Force in Newton (GrF) using a Leonardo Mechanograph® GF myometer.

Inclusion and exclusion criteria: Inclusion criteria were defined as following: (1) children and adolescents equal to or greater 4 years up to 17 years of age, (2) absence of any illness or disturbance of the locomotor system, (3) ability to cooperate during analysis, (4) written consent of caregivers (and children if age of 6 years or older). Exclusion criteria included declaration of pain during walking or climbing of stairs and presence of any other physical or neurologic disease.

Recruitment of patients: Eligible healthy children and adolescents were recruited in different regional kindergartens and in elementary and secondary schools. All parents received short questionnaires inquiring regular physical activities, the time per week consuming media of all kind and the potential presence of impairments affecting the ability to climb stairs. All healthy participants were assessed only once.

Protocol: All mechanographic assessments were performed following a standardized protocol with defined order of assessments, duration of breaks between assessments and determined verbal explanation of assessments (see Table S2). In brief, participants stood with both feet on the lower plateau of the device and were asked to climb the four stairs as quickly as possible without

Table 1. Nomenclature of mechanographic outcome parameters.

Name	SI unit	Explanation
Stair climb test (SC up and SC down)		
• *av.v.hor*	[m/s]	average vertical velocity, which is the projection of the speed to the horizontal plain
• *av.F.max.rel*	[Fg]	average over all analyzed steps of maximum force per step as multiples of body weight (Fg)
• *av.F.min.rel*	[Fg]	average over all analyzed steps of minimum force per step as multiples of body weight (Fg)
• *av.P.rel*	[W/kg]	Average power over all analyzed steps in relation to body mass
Chair rising test (CRT)		
• *av.P.max.rel*	[W/kg]	Mean maximal power during rising phase in relation to body mass
Grip force myometry		
• *F.max*	[N]	Maximal force during hand grip in Newton

Table 2. Baseline auxology of healthy probands with indication of mean mass. Height and BMI with standard deviations according to age and sex.

	All				Male				Female			
Age [ys]	n=	Mass [kg]	Height [cm]	BMI [kg/m²]	n=	Mass [kg]	Height [cm]	BMI [kg/m²]	n=	Mass [kg]	Height [cm]	BMI [kg/m²]
4	17	16.7±1.6	104.1±4.6	15.4±1.2	6	17.1±1.8	103.3±2.4	15.2±1.5	11	16.5±1.5	102.8±5.1	15.6±1.1
5	11	20.3±2.5	114.2±6.1	15.6±1.3	6	19.8±1.85	113.7±5.5	15.3±1.0	5	20.9±3.3	114.8±7.3	15.8±1.6
6	25	23.8±3.6	119.6±4.9	16.6±1.7	14	24.2 ± 3.9	119.2±4.4	17.0±1.8	11	23.3±3.1	120.0±5.6	16.1±1.4
7	25	26.6±4.8	126.7±7.4	16.5±2.0	13	27.9±5.8	126.5±9.1	17.3±2.1	12	25.2±3.0	126.9±5.5	15.7±1.6
8	23	29.9±4.8	132.8±5.3	16.9±2.4	12	29.7±5.9	131.5±4.4	17.1±2.7	11	30.2±3.6	134.3±6.1	16.8±2.1
9	39	33.6±6.6	138.2±5.0	17.5±3.0	24	34.3±7.1	138.6±4.9	17.8±3.0	15	32.4±5.6	137.5±5.1	17.2±2.9
10	25	34.2±6.4	141.4±6.7	17.0±2.4	13	35.2±7.7	142.3±7.7	17.3±2.6	12	33.0±4.8	140.5±5.7	16.7±2.1
11	39	45.2±13.3	152.1±8.4	19.3±4.1	21	43.2±11.2	149.6±7.3	19.1±3.5	18	47.5±15.3	155.0±8.9	19.4±4.9
12	25	45.0±10.2	154.3±7.0	18.8±3.8	14	45.7±12.0	153.1±8.1	19.4±4.3	11	44.0±7.7	155.8±5.3	18.1±2.9
13	19	57.2±15.5	161.2±7.5	21.9±5.1	9	60.1±20.4	161.3±9.0	22.8±6.5	10	54.7±9.9	161.0±6.3	21.0±3.7
14	26	54.4±10.1	166.8±8.9	19.4±2.5	9	57.4±12.9	170.8± 9.2	19.5±2.8	17	52.8±8.3	164.8±8.2	19.4±2.4
15	8	57.0±6.2	167.1±8.4	20.5±1.8	2	54.5±14.0	169.0±19.8	18.9±0.4	6	57.9±3.3	166.5±54.4	21.0±1.8
16	6	63.3±7.1	171.1±7.1	21.6±2.6	2	67.4±3.6	178.5±7.8	21.3±3.0	4	61.2±7.9	167.5±3.1	21.8±2.8

running and then stand still on the upper plateau. Use of the handrail was allowed if necessary. The step-down maneuver was performed accordingly. Sequence of assessments was (1) 4SC-Up and 4SC-Dn, (2) GrF, (3) CRT. All assessments were performed three times after standardized breaks of 30 seconds. The times to perform 4SC-Up, 4SC-Dn and CRT were also measured manually.

The Leonardo Mechanography software analyses the stair ascent and descents and segments the continuous movement into individual steps based on analysis of ground reaction force variation as well as the variation of the position (Center of Force, CoF, often also referred to as Center of Pressure, CoP) in relation to the individual stairs step position. For each detected step, the maximum and minimum force as well as the average velocity and peak power are analyzed and average values over all analyzed steps are calculated. As established for other mechanographic for inter-individual comparison and cross-sectional reference data force values are normalized to body weight while power values are normalized to body mass[24–26]. The main mechanographic outcome parameters or the performed tests are summarized in Table 1.

Statistical analysis

We analyzed clinical data descriptively and processed them with absolute frequencies and percentage values. Descriptive statistics were calculated for the entire study population and are presented for sex and age. For statistical analysis SPSS (version 26.0) was used. We used parametric methods for normally distributed data and non-parametric methods (median and 25/75% quartile) for other data. For correlation analysis, Kendall-Tau-b coefficient was calculated. A p-value lower 0.05 was considered statistically significant for all analyses. Percentiles were calculated using the LMS method provided by the GAMLSS package in R (v4.1.0, The R Project for Statistical Computing; method BCPE) with following parameters: Mu=2.0; Sigma=1.0; Nu=0.5. A detailed

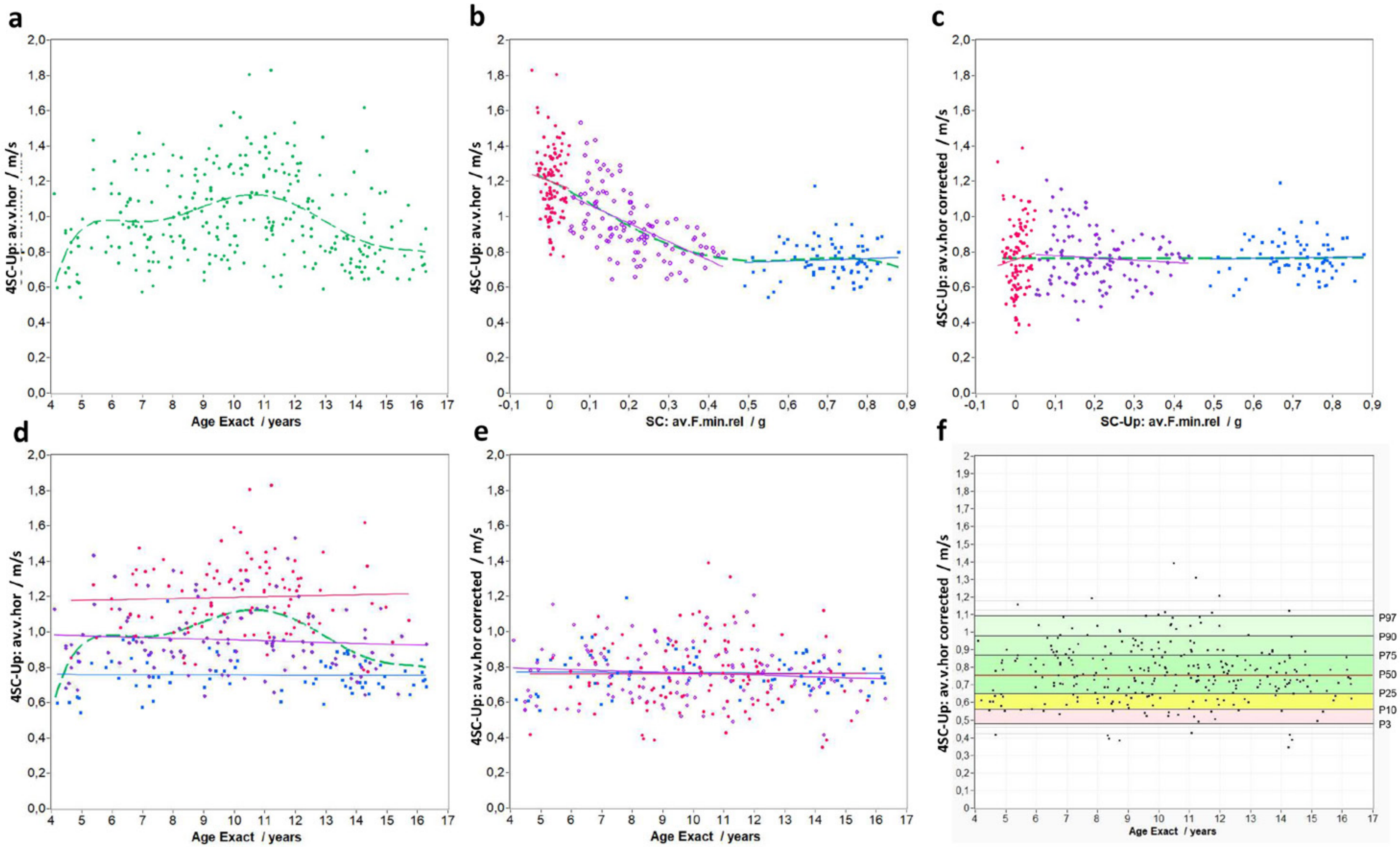

Figure 1. 4SC-Up data for av.v.hor: a) vs age, b) vs. *av.F.min.rel* including sub-groups of slow walkers (blue squares), fast walkers (purple hollow circles) and runners (red circles) c) after compensation of group differences d) vs. age including linear interpolation per sub-group e) after compensation f) resulting age-independent reference data (percentiles).

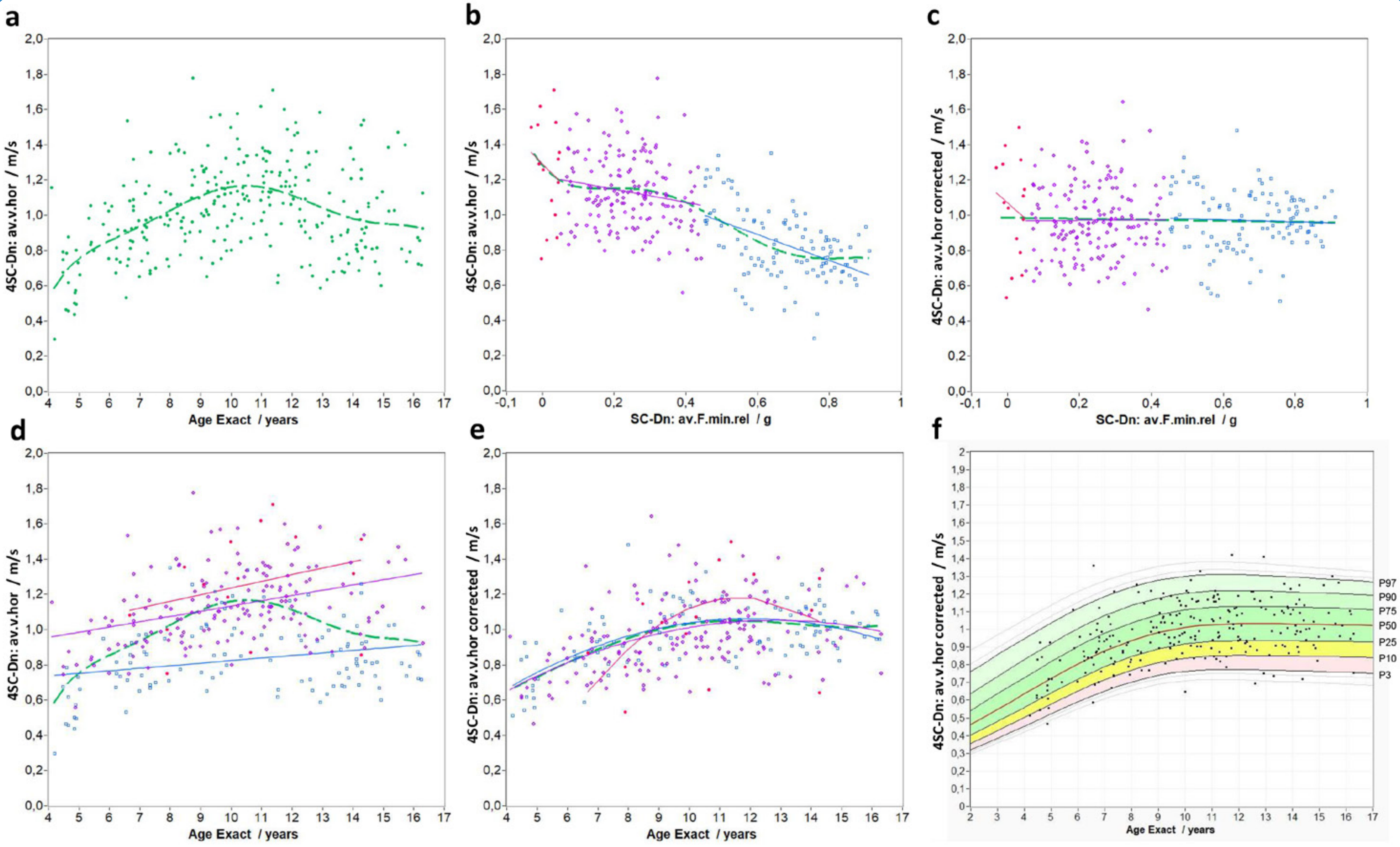

Figure 2. 4SC-Dn data for *av.v.hor*: a) vs age, b) vs. *av.F.min.rel* including sub-groups of slow walkers (blue squares), fast walkers (purple hollow circles) and runners (red circles) c) after compensation of group differences d) vs. age including linear interpolation per sub-group e) after compensation f) resulting age-independent reference data (percentiles).

description of the used methods can be found elsewhere[26]. Based on these percentiles and the resulting LMS parameters z-Scores were calculated according to:

$((y/M)L-1)/S*L$ for $L \neq 0$

$1/S*\ln(y/M)$ for $L=0$

Formular 1

L,M,S: Parameters per age group calculated by LMS method, y: individual measurement result

Results

A cohort of 301 otherwise healthy children and adolescents aged 4 to 16 years underwent mechanographic assessments in two kindergartens, two elementary school and two high schools (one comprehensive school, one academic high school) between July 2019 and March 2020 (150 females and 151 males). Data of one 4-year-old boy were excluded, as he did not want to participate during the appointed measurement.

Due to technical problems, mechanographic analyses of the SC test were not available in 12 individuals who were therefore excluded, leaving data of 288 participants (144 males and 144 females) for analysis. Grip force results for the dominant hand were available for all 288 subjects and CRT was available for 287 children (one 4-year-old girl refused participation after completion of the previous assessments). The baseline characteristics of all participants are displayed in Table 2. All measurements were performed without occurrence of any unexpected incidents like falls or injuries. The ethnicity of subjects was predominantly middle European (72.6%), followed by mixed middle European (16.7%), Caucasian (3.1%) and Arabic-Mediterranean (3.1%) and other ethnicities (4.5%).

Questionnaire

We received completed questionnaires for 285 participants (99.0%). Regular physical activity was reported for 266 individuals (92.4%) with a mean of 2.1±0.5 hours/week (min/max: 1.0; 4.0). Estimated screen-time as a parameter of media consumption was available for 281 participants (97,6%) with a mean of 6.3±5.6 hours/week (min/max: 0;30.0). Presence of physical disabilities was reported for 11 individuals (hearing problems 1; tremor of hands 1, visual disorder 2, non-motor development issues 3, migraine 1, heart condition 1, no further information available 2).

Four step stair ascent and descent (4SC-Up and 4SC-Dn)

Mechanographic analyses - compensation of movement strategies

First analysis of the *av.v.hor* data revealed that there was no significant gender difference but an unexpected age dependency of the data (Figure 1a) with a peak in *av.v.hor* in the age group between 9 and 12 years. As this finding contradicts other studies reporting age dependencies[11,24] we further analyzed the individual measurements manually. These analyses revealed that despite the command for the stair climb being not to run, nevertheless a variety of movement strategies were used. Movement strategies ranged from slow walking over fast walking to almost running. An effective parameter to separate these movement strategies was found to be the average minimum force per step *av.F.min.rel*. Since in opposite to walking where always one foot is in contact with the ground, in running there is a flight phase without ground contact and hence a minimum ground reaction force of zero. Faster running typically results in longer flight times and periods of a ground reaction force of zero. Figure 1b underlines this assumption where below a threshold value of 0.5**Fg* a linear increase of *av.v.hor* with decreasing *av.F.min.rel* can be observed. These three clusters become even more visible when scatter plots of all three measurements per participant are analyzed, as can be seen in the supplement Figure S2.

Figure 1b also shows the separation of three groups based on the following *av.F.min.rel* thresholds derived from the three clusters observed in this scatter plot:

- Slow Walkers: *av.F.min.rel* >0.45**Fg* (blue boxes)
- Fast Walkers: 0.05 *Fg* >*av.F.min.rel* > 0.45**Fg* (purple circles)
- Runners: *av.F.min.rel* < 0.05**Fg* (red dots)

Fg being the body weight (the ground reaction force of the body mass due to earth's gravity, 1g = 9.81 m/s^2)

When grouping *av.v.hor* values according to these three groups Figure 1c shows that each of the three subgroups shows values widely independent of age. The apparent age dependency in the overall group can be explained by a dominant proportion of runners in the age group between 9 and 12 compared to the other age groups where the proportion was more or less equally distributed to all three groups. Additional information regarding the established algorithm is available in the supplemental data.

After this compensation for 4SC-Up no age dependency could be observed any more (Figure 1e). Based on this corrected data LMS percentiles were calculated. These data were further used to calculate z-Scores for different patient groups (data will be reported separately). The same method was used for normalization of the parameters *av.P.rel* and *av.F.max.rel*. Figure 2a-f show this analysis concept applied to 4SC-Dn results. Other than in the 4SC-Up test, all parameters of the 4SC-Dn test showed a distinct age dependency even after applying the described normalization algorithm.

Time measurement (manually vs. automated) and selection of test results

Age adjusted percentiles for the required time to climb 4 standardized steps are displayed in Figure 3. Manually measured time by stopwatch and device time ('time analysed' [sec]) of the 4SC differed significantly in the overall cohort with shorter times measured by the device (measurement with best manually tested time; mean times 1.27 vs. 1.46 sec., mean difference -0.19 sec. (CI:-0.23;-0.15); p<0.001). Respective differences were evident in almost all age cohorts. Time results according to age and sex are displayed in Table 3.

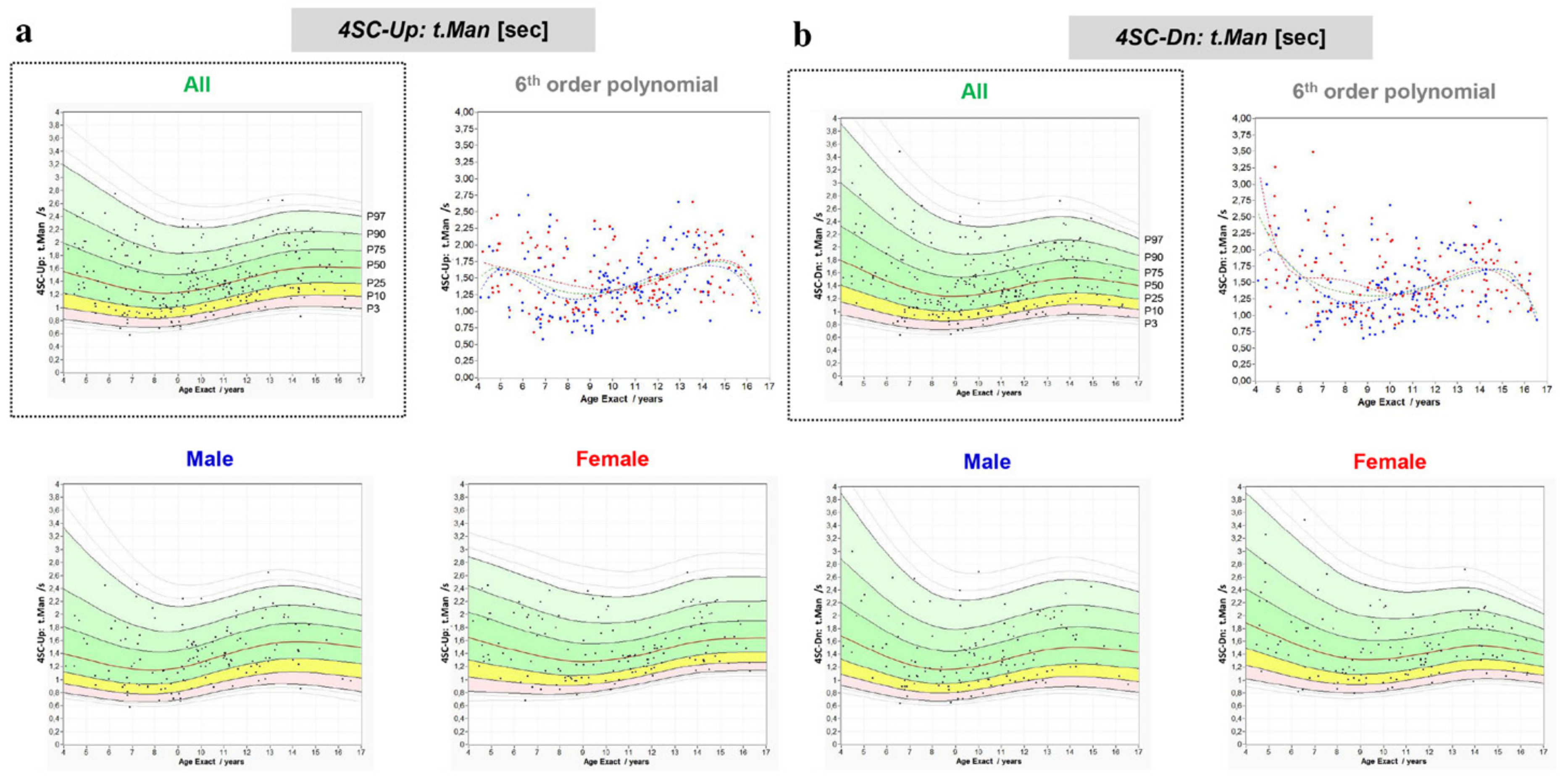

Figure 3. Age-adjusted percentiles of manually measured time required to climb (left) and ascend (right) 4 stairs. Percentiles are shown for both sexes (upper left), separate sexes (below). Sex differences appear marginal - as displayed by 6th order polynomial with interpolation (upper right).

Since multiple measurements were recorded per individual, in accordance to published reference data[24] a selection of the best measurement which is then used for analysis is needed. To minimize potential influence by operators on outcome parameters in clinical application, instead of using a manual selection based on the minimum manually recorded measured time (reflecting to some extent the stair climb performance in terms of maximum speed), an automated selection method was implemented based on the maximum *av.h.hor corrected*. We compared all presented outcome parameters using either manual or automated selection. Using the automated selection instead of the manual selection the different outcome parameters showed results indicating an improved performance (e.g. shorter time or

Table 3. Manual measured and device measured mean time required to ascend 4 stairs (4SC-Up) according to age and sex groups. Results refer to the measurement with the lowest manually measured time out of 3 performed measurements.

	All			Male			Female		
Age [ys]	n=	manual time [sec]	device time [sec]	n=	manual time [sec]	device time [sec]	n=	manual time [sec]	device time [sec]
4	17	1.65 ±0.41	1.45 ±0.36	6	1.48 ±0.30	1.33 ±0.34	11	1.74 ±0.44	1.52 ±0.38
5	11	1.34 ±0.36	1.27 ±0.22	6	1.24 ±0.34	1.25 ±0.19	5	1.46 ±0.37	1.30 ±0.27
6	25	1.47 ±0.65	1.25 ±0.37	14	1.45 ±0.75	1.12 ±0.24	11	1.51 ±0.55	1.42 ±0.44
7	25	1.55 ±0.61	1.25 ±0.26	13	1.52 ±0.63	1.17 ±0.17	12	1.59 ±0.61	1.35 ±0.32
8	23	0.99 ±0.17	1.15 ±0.19	12	0.93 ±0.16	1.11 ±0.22	11	1.06 ±0.17	1.19 ±0.15
9	39	1.39 ±0.42	1.21 ±0.26	24	1.20 ±0.37	1.16 ±0.20	15	1.53 ±0.48	1.30 ±0.33
10	25	1.36 ±0.39	1.11 ±0.22	13	1.33 ±0.37	1.09±0.21	12	1.39 ±0.41	1.13 ±0.24
11	39	1.40 ±0.26	1.15 ±0.26	21	1.45 ±0.29	1.19 ±0.33	18	1.34 ±0.22	1.11 ±0.15
12	25	1.52 ±0.44	1.24 ±0.31	14	1.51 ±0.47	1.25 ±0.34	11	1.52 ±0.42	1.24 ±0.28
13	19	1.79 ±0.35	1.54 ±0.35	9	1.85 ±0.32	1.59 ±0.33	10	1.73 ±0.38	1.49 ±0.38
14	26	1.67 ±0.39	1.45 ±0.31	9	1.54 ±0.39	1.30 ±0.28	17	1.73 ±0.38	1.52 ±0.30
15	8	1.61 ±0.21	1.45 ±0.29	2	1.54 ±010	1.21 ±0.09	6	1.64 ±0.23	1.53 ±0.29
16	6	1.40 ±0.31	1.55 ±0.27	2	1.36 ±0.49	1.77 ±0.10	4	1.42 ±0.30	1.45 ±0.27

Table 4. Selected Results of mechanographic analysis during the 4 stair ascent (4SC-Up) according to age and sex. Median results and quartiles (P25;P75) are displayed for all parameters.

	All				Male				Female			
Age [ys]	n=	Av.P.rel [W/kg]	Av.F.max. rel [Fg]	Av.V.hor [m/s]]	n=	Av.P.rel [W/kg]	Av.F.max. rel [Fg]	Av.V.hor [m/s]]	n=	Av.P.rel [W/kg]	Av.F.max.rel [Fg]	Av.V.hor [m/s]]
4	17	3.6 (3.1;4.3)	1.4 1.3;1.6)	0.6 (0.6;0.7)	6	4.4 (4.2;4.8)	1.4 (1.2;1.5)	0.8 (0.6;0.9)	11	3.3 (3.1;3.6)	1.4 (1.3;1.7)	0.6 (0.6;0.6)
5	11	3.7 (3.5;4.2)	1.4 (1.4;1.7)	0.8 (0.7;0.9)	6	3.8 (3.5;4.2)	1.4 (1.3;1.7)	0.8 (0.8;1.0)	5	3.6 (3.5;4.0)	1.4 (1.3;1.6)	0.8 (0.6;0.9)
6	25	4.2 (3.6;4.6)	1.4 (1.3;1.5)	0.8 (0.7;0.9)	14	4.3 (3.6;4.7)	1.4 (1.3;1.6)	0.8 (0.7;0.9)	11	4.0 (3.6;4.4)	1.4 (1.3;1.5)	0.8 (0.7;0.9)
7	25	3.8 (3.2;4.2)	1.5 (1.4;1.5)	0.8 (0.7;0.9)	13	3.9 (3.4;4.3)	1.4 (1.4;1.5)	0.8 (0.7;0.9)	12	3.6 (3.1;4.0)	1.5 (1.5;1.6)	0.8 (0.7;1.0)
8	23	3.2 (2.9;3.5)	1.4 (1.4;1.6)	0.7 (0.6;0.8)	12	3.3 (2.9;3.9)	1.4 (1.4;1.6)	0.7 (0.6;0.9)	11	3.3 (2.9;3.5)	1.4 (1.4;1.6)	0.6 (0.6;0.7)
9	39	3.6 (3.1;4.0)	1.4 (1.4;1.5)	0.8 (0.7;0.9)	24	3.7 (3.5;4.3)	1.4 (1.4;1.5)	0.8 (0.7;0.9)	15	3.5 (3.1;3.9)	1.5 (1.3;1.5)	0.7 (0.6;0.8)
10	25	3.8 (3.4;4.3)	1.4 (1.4;1.5)	0.8 (0.7;1.0)	13	4.0 (3.1;4.7)	1.4 (1.4;1.5)	0.8 (0.7;1.0)	12	3.8 (3.6;4.2)	1.5 (1.4;1.5)	0.8 (0.7;1.0)
11	39	3.7 (3.2;4.2)	1.4 (1.3;1.5)	0.7 (0.6;0.9)	21	3.9 (3.2;4.2)	1.4 (1.4;1.5)	0.8 (0.6;1.0)	18	3.7 (3.3;3.9)	1.5 (1.3;1.5)	0.7 (0.7;0.8)
12	25	3.6 (3.3;3.9)	1.5 (1.4;1.6)	0.8 (0.7;0.8)	14	3.7 (3.4;4.1)	1.4 (1.4;1.5)	0.8 (0.7;0.8)	11	3.3 (3.1;3.6)	1.5 (1.5;1.7)	0.7 (0.6;0.8)
13	19	3.7 (3.3;4.1)	1.4 (1.3;1.5)	0.7 (0.7;0.8)	9	3.9 (3.2;4.3)	1.4 (1.3;1.5)	0.7 (0.7;0.8)	10	3.6 (3.5;4.2)	1.4 (1.3;1.5)	0.8 (0.7;0.9)
14	26	3.8 (3.3;4.1)	1.4 (1.3;1.5)	0.7 (0.7;0.8)	9	3.8 (3.1;4.2)	1.5 (1.3;1.5)	0.8 (0.6;0.9)	17	3.7 (3.3;4.2)	1.4 (1.3;1.6)	0.7 (0.7;0.8)
15	8	3.6 (3.0;3.9)	1.4 (1.4;1.5)	0.7 (0.6;0.7)	2	4.0 (4.0;-)	1.4 (1.3;-)	0.7 (0.7;-	6	3.3 (2.9;3.8)	1.4 (1.4;1.5)	0.6 (0.5;0.7)
16	6	3.7 (3.3;4.2)	1.5 (1.4;1.6)	0.7 (0.7;0.8)	2	3.5 (3.5;-)	1.5 (1.4;-)	0.8 (0.7;-)	4	4.0 (2.9;4.3)	1.4 (1.4;1.5)	0.7 (0.6;0.8)

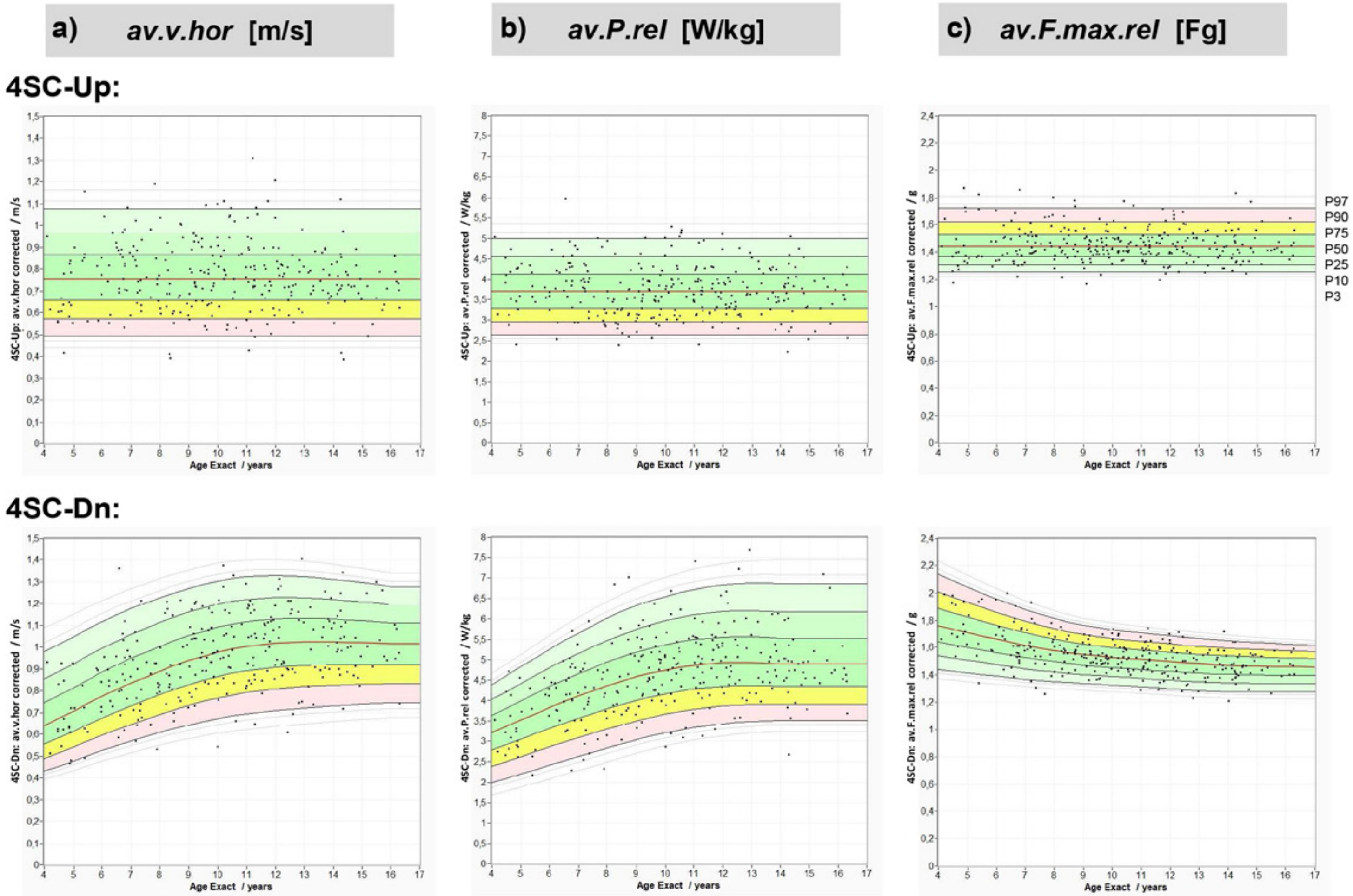

Figure 4. Scatter plots and age-dependent LMS percentiles for selected outcome parameters, all compensated for *av.F.min.rel* = 0.5**Fg*. Upper plots: stair ascent (4SC-Up), lower plots: stair descent (4SC-Dn).

higher power) between 1% and 9%. At the same time a decreased variability over all individuals between 0.1% and 4% depending on the parameter was found. Hence out of the three performed measurements per individual, the highest av.v.hor corrected was selected for further analysis. All following parameters are corrected for *av.F.min.rel* = 0.5**Fg*.

av.P.rel corrected [W/kg]

Figure 4a shows the results for av.P.rel corrected. For 4SC-Up tests neither age nor gender dependency was found. For 4SC-Dn a relevant age dependency could be observed with increasing values up until the age of 12 years. Results of the average (relative) power (*av.P.rel* [W/kg]) are shown in Table 4 separated for male and female participants in relation to age categories. Results were comparable for both sexes in children aged 9 years and older, showing a peak in children aged 8 to 11. Figure 4b displays the calculated percentiles for 4SC-Up and 4SC-Dn according to age categories on basis of *av.P.rel* results. Percentiles according to mass, height and BMI of subjects are displayed in the appendix (Figure S3). After the correction described above there was no significant correlation between the main outcome parameters and the manual assessed total measurement time apart from *F.min.rel* (Figure S4).

Correlation analysis showed no apparent association of corrected *av.P.rel* with height, mass, BMI or with questionnaire data regarding hours of screen time, but a significant correlation with questionnaire data regarding hours of physical exercise per week (Kendall-Tau-b: 0.13, $p=0.009$).

We observed a fair test-retest reliability between the first and the second measurement of *av.P.rel* for 4SC-Up (ICC all: 0.82; males: 0.75; females: 0.89) and between the first and the third measurement (ICC all: 0.78; males: 0.67; females: 0.90).

av.F.max.rel corrected [Fg] and av.v.hor.corrected [m/s]

Figure 4a and c show the results for both parameters. Neither age nor gender dependency were found. A relevant age dependency could be observed in both parameters for 4SC-Dn up prior to age 9 to 12. Percentiles of *av.F.max.rel* corrected and *av.v.hor* corrected according to weight, height and BMI of participants are displayed in the appendix (Figure S4 and S5).

Correlation with grip force (GF) and chair rising test (CRT)

Results of GF and CRT measurements are displayed in Table S3. As expected, the maximal force during the hand-grip manoeuvre (F.max) in Newton correlated significantly with the average power during the ascent of 4 steps (Kendall-Tau-b: $r=0.55$, $p<0.001$). However, there was no correlation with hand grip results when the average power during the four stair climb was adjusted to body mass (*av.P.rel*, relative power). There was mild but significant correlation between *P.max.rel* of the CRT manoeuvre and *av.P.rel* (Kendall-Tau-b: $r=0.14$, $p<0.001$) during the ascent of 4 stairs.

Discussion

Different aspects are important to consider, when mechanographic results are interpreted:

(i) The analysis of the recorded absolute values of force and power is difficult in cross-section. Rather, these results have to be related to the individual body weight or body mass respectively[24]. (ii) The interpretation of the used force has to be reconsidered. The force expended for a movement is very dependent on the movement strategy and a higher force not necessarily results in an increased movement outcome (e.g. *av.v.hor*). This becomes particularly obvious in the 4SC-Dn data, where both, the horizontal velocity as well as the relative power increase with age until about age 12 while the maximum force decreases in the same age span. These observations are in line with age-related effects observed in other mechanographic studies[11,24,28], and may be explained by increased function and improved coordination during growth.

Interestingly, these aging effects were not evident in the 4SC-Up data. One explanation might be that 4SC-Up and 4SC-Dn tests actually address two fundamentally different aspects of muscle function. Both rely on very different demands on muscle function as well as on coordination and therefore most likely reflect different influences on the measurement parameters by growth. While 4SC-Up is focussing on generation of energy (creating potential energy by lifting the centre of mass of the body to a certain height), 4SC-Dn is focussing on dissipation of energy[27,28] (or more physically correct: converting potential energy to heat). For 4SC-Dn decelerating (breaking the fall) is the major task for muscle function. However, an efficient deceleration process implies continuity: Rather than short intervals of hard breaking resulting in high force peaks for a short time, a more continuous movement appears to be more effective, allowing lower peak forces and a continuous deceleration/energy dissipation process.

An explanation for the observed differences in age-dependency might therefore be that (i) the stair descent from the coordination aspect is the more challenging task and (ii) the general muscle function of generating energy (stair ascent) might be the more commonly used in every-day living and therefore better trained compared to stair descent. In this context, we believe that both movement tasks (4SC-Up as well as 4SC-Dn) should be assessed since they target two very different aspects of muscle function[26,28].

Apart from the mechanographic results, we found the time measured by the device to be significantly shorter than the corresponding manual measurement of the same test. This difference apparently reflects the delay between the "Go" of the examiner and the actual initiation of the motion sequence of the individual in conventional execution, whereas the mechanograph uses its variation of force and of position to segment steps and based on this segmentation calculates the time from start to the end of the movement cycle – reaction latencies are irrelevant for this method. The duration calculated by the device can therefore be regarded as more accurate than the manually measured times. The mean difference of 0.19 seconds constitutes ~15% of the measured absolute values – suggesting a substantial measurement uncertainty in manual testing, assuming that this finding is reproducible in other studies and larger cohorts. In addition, comparison of the two methods also showed a lower variance of the time calculated by the device (0.27s) compared to the manual time measurement (1.46s). The 4SC was performed first and bias by exhaustion can be excluded. This might explain that our results were lower than in the available reference data in boys aged 4 and 5 years where the 4SC was performed as one of other timed function tests[18].

Our data show only little and no consistent sex differences for both, manually measured time of the 4SC and corresponding mechanographic parameters. This is quite unexpected since data of other mechanographic studies in young and healthy subjects as the chair rising test exhibit sex differences[11], where boys also achieved the higher performance values even when corrected for body mass in both, our study and published reference data[24]. On the other hand, we found that movement strategy while climbing stairs ('Runner' vs. 'Walker') clearly influences the mechanographic results. Although the standardized commands clearly asked participants to walk and not to run the 4 steps as fast as possible, the applied movement strategy needs to be controlled more thoroughly in future studies. On the other hand, in quite fit subjects not allowing maximum performance (like running) is an artificial limitation when targeting a maximum outcome. As can be seen in our data, this obviously is subject to individual decisions and hence difficult to normalize. However, this effect is less likely to be of relevance when addressing patient groups with limited performance. Nevertheless, in this context we believe the proposed normalization to *av.F.max.rel* = 0.5**Fg* is a very feasible and at the same time clinically practical compromise to compensate this systematic limitation of stair climb tests in general.

We combined mechanographic analysis of the 4-step stair ascent and descent with yet established other apparatus-based tests like the mechanographic analysis of the CRT and myometric measurement of maximal hand grip force. Summarized, the data gathered in our study corresponded with published reference data of CRT[24] and GF[24,27,28].

Though not evaluated in detail, we experienced motivation and collaboration of participants to be very good. The stair-bound mechanographic measurements were easy and quick to understand and perform. The mechanographic measurement requires short periods of standing still before and after ascent. Apart from preschool children, where first measurements were sometimes erroneous due to a lack of understanding of the correct movement task, the gathered data was predominantly valid and interpretable.

Our study has some obvious limitations: First, our cohort was not big enough to analyze data separately for different ethnic groups; and second, we did not carry out a qualitative assessment of climbing stairs[28]. While virtually all children aged 6 years and older climbed stairs in an alternating mode, we experienced expectable differences in younger participants. Third, the composition of our cohort was rather heterogeneous with fewer participants in older age cohorts, so that the calculated percentiles for adolescents aged 15 and older have to be interpreted with caution. Fourth, all participants were advised to climb the stairs as quick as possible without running. However, we identified different movement strategies that were used by participating children and adolescents and thus complicated comparison of results. We are aware that this is an methodological issue. Nevertheless, we believe that a fully standardized implementation in a pediatric cohort might be difficult to impossible. The presented algorithm therefore represents a defensible analysis option for comparative analysis across all motion strategies.

In conclusion, we present reference data for the timed-function test 4SC for stair ascent as well as stair descent in healthy children and adolescents and mechanographic reference data for the respective motion sequence. We present a simple algorithm to compensate effects caused by different movement strategies. After compensation a clear difference of the main outcome parameters stair ascent and descent regarding age dependency is obvious. We therefore believe, that assessing both movement tasks is preferable, since the two tasks address very different aspects of muscle function.

We believe mechanographic analysis of stair-climbing to be a promising tool for evaluation of muscle strength and function of the lower extremities as it enables physically exact measurements of a highly relevant activity of daily living.

Authors' contributions

David Christof Schorling: Conception of the study, writing of study protocol, recruitment of participants, conduct of measurements, analysis and interpretation of data, writing of first manuscript. Rainer Rawer: Consultancy in technical issues/analysis. Imke Kuhlmann: Recruitment of participants, conduct of measurements, analysis of data, revising of manuscript. Cornelia Mülller: Conduct of measurements, revising of manuscript. Astrid Pechmann: Interpretation of data, revising of manuscript. Janbernd Kirschner: Conception of the study, revising study protocol, analysis and interpretation of data, revising of manuscript.

Funding

This study was supported by intramural research funding from the Department of Neuropediatrics and Muscle Disorders. Novotec GmbH provided technical support (RW) but did not provide any additional material or financial support for the conduct of this study.

Acknowledgements

We thank all participants of this study and their parents.

References

1. O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscul Disord NMD 2007;17(9-10):693-697.
2. Scott E, Eagle M, Main M, Sheehan J. The North Star Ambulatory Assessment. Published online January 2006.
3. Florence JM, Pandya S, King WM, et al. Intrarater reliability of manual muscle test (Medical Research Council scale) grades in Duchenne's muscular dystrophy. Phys Ther 1992;72(2):115-122; discussion 122-126.
4. Dupépé EB, Davis M, Elsayed GA, et al. Inter-rater reliability of the modified Medical Research Council scale in patients with chronic incomplete spinal cord

injury. J Neurosurg Spine. Published online January 18, 2019:1-5.
5. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166(1):111-117.
6. Andriacchi TP, Andersson GB, Fermier RW, Stern D, Galante JO. A study of lower-limb mechanics during stair-climbing. J Bone Joint Surg Am 1980;62(5):749-757.
7. McFadyen BJ, Winter DA. An integrated biomechanical analysis of normal stair ascent and descent. J Biomech 1988;21(9):733-744.
8. Nightingale EJ, Pourkazemi F, Hiller CE. Systematic review of timed stair tests. J Rehabil Res Dev 2014;51(3):335-350.
9. Riener R, Rabuffetti M, Frigo C. Stair ascent and descent at different inclinations. Gait Posture 2002;15(1):32-44.
10. Ni M, Brown LG, Lawler D, Bean JF. Reliability, Validity, and Minimal Detectable Change of Four-Step Stair Climb Power Test in Community-Dwelling Older Adults. Phys Ther 2017;97(7):767-773.
11. Zaino CA, Marchese VG, Westcott SL. Timed up and down stairs test: preliminary reliability and validity of a new measure of functional mobility. Pediatr Phys Ther Off Publ Sect Pediatr Am Phys Ther Assoc 2004;16(2):90-98.
12. McDonald CM, Henricson EK, Abresch RT, et al. The cooperative international neuromuscular research group Duchenne natural history study - a longitudinal investigation in the era of glucocorticoid therapy: design of protocol and the methods used. Muscle Nerve 2013;48(1):32-54.
13. Smith EC, Conklin LS, Hoffman EP, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy: An 18-month interim analysis of a non-randomized open-label extension study. PLoS Med 2020;17(9):e1003222.
14. Bushby K, Finkel R, Wong B, et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. Muscle Nerve 2014;50(4):477-487.
15. Victor RG, Sweeney HL, Finkel R, et al. A phase 3 randomized placebo-controlled trial of tadalafil for Duchenne muscular dystrophy. Neurology 2017; 89(17):1811-1820.
16. McDonald CM, Campbell C, Torricelli RE, et al. Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Lond Engl 2017;390(10101):1489-1498.
17. Goemans N, Mercuri E, Belousova E, et al. A randomized placebo-controlled phase 3 trial of an antisense oligonucleotide, drisapersen, in Duchenne muscular dystrophy. Neuromuscul Disord NMD 2018;28(1):4-15.
18. Hoskens J, Goemans N, Feys H, De Waele L, Van den Hauwe M, Klingels K. Normative data and percentile curves for the three-minute walk test and timed function tests in healthy Caucasian boys from 2.5 up to 6 years old. Neuromuscul Disord NMD 2019;29(8):585-600.
19. McGrath R. Maximal Handgrip Strength Alone Could Be an Incomplete Measure of Muscle Function. J Am Med Dir Assoc 2021;22(4):882-883.
20. Fricke O, Weidler J, Tutlewski B, Schoenau E. Mechanography - a new device for the assessment of muscle function in pediatrics. Pediatr Res 2006; 59(1):46-49.
21. Sumnik Z, Matyskova J, Hlavka Z, Durdilova L, Soucek O, Zemkova D. Reference data for jumping mechanography in healthy children and adolescents aged 6-18 years. J Musculoskelet Neuronal Interact 2013;13(3):297-311.
22. Matheson LA, Duffy S, Maroof A, Gibbons R, Duffy C, Roth J. Intra- and inter-rater reliability of jumping mechanography muscle function assessments. J Musculoskelet Neuronal Interact 2013;13(4):480-486.
23. Bendixen RM, Lott DJ, Senesac C, Mathur S, Vandenborne K. Participation in daily life activities and its relationship to strength and functional measures in boys with Duchenne muscular dystrophy. Disabil Rehabil 2014;36(22):1918-1923.
24. Busche P, Rawer R, Rakhimi N, Lang I, Martin DD. Mechanography in childhood: references for force and power in counter movement jumps and chair rising tests. J Musculoskelet Neuronal Interact 2013;13(2):213-226.
25. Saxer S, Speich R, Toigo M, Mueller SM, Ulrich Somaini S. Reliability of parameters during stair ascent measured with Leonardo Mechanograph(®) Stair A in healthy subjects. J Musculoskelet Neuronal Interact 2015;15(3):257-263.
26. Veilleux LN, Robert M, Ballaz L, Lemay M, Rauch F. Gait analysis using a force-measuring gangway: intrasession repeatability in healthy adults. J Musculoskelet Neuronal Interact 2011;11(1):27-33.
27. Beenakker EAC, van der Hoeven JH, Fock JM, Maurits NM. Reference values of maximum isometric muscle force obtained in 270 children aged 4–16 years by hand-held dynamometry. Neuromuscul Disord 2001;11(5):441-446.
28. Fernandes LAY, Caromano FA, Hukuda ME, Escorcio R, Carvalho EV. Elaboration and reliability of functional evaluation on going up and downstairs scale for Duchenne Muscular Dystrophy. Rev Bras Fisioter Sao Carlos Sao Paulo Braz 2010;14(6):518-526.

Supplemental data

Methods

Outcome measures

Further information regarding the Leonardo Stair Mechanogaph:

The Leonardo Mechanograph® Stair system uses one sensor in each corner of the device and thus measures the total force acting on the complete stair case instead of the ground reaction force per foot (as typically used when presenting ground reaction forces related to gait cycles). Instead, this set-up allows the additional analysis of the trajectory of force vector entering the surface of the stair case (center of force, *CoF*, often also referred to as center of pressure, *CoP*). The measurement concept has been described for a similar Mechanograph system assessing gait on a horizontal surface[26] and reproducibility data for basic outcome parameters using as similar staircase (Leonardo Mechanograph® Stair A, 5 steps, 18cm step height) for adults 25 have been reported elsewhere.

Results

Four step stair ascent and descent (4SC-Up and 4SC-Dn)

Mechanographic analyses - compensation of movement strategies – more detailed information

First analysis of the av.v.hor data revealed that there was no significant gender difference but an unexpected age dependency of the data (Figure 1a) with a peak in *av.v.hor* in the age group between 9 and 12 years. As this finding contradicts other studies reporting age dependencies[11,24] we further analyzed the individual measurements manually. These analyses revealed that despite the command for the stair climb being not to run, nevertheless a variety of movement strategies were used. Movement strategies ranged from slow walking over fast walking to almost running. An effective parameter to separate these movement strategies was found to be the average minimum force per step *av.F.min.rel*. Since in opposite to walking where always one foot is in contact with the ground, in running there is a flight phase without ground contact and hence a minimum ground reaction force of zero. Faster running typically results in longer flight times and periods of a ground reaction force of zero. Figure 1b underlines this assumption where below a threshold value of 0.5**Fg* a linear increase of *av.v.hor* with decreasing *av.F.min.rel* can be observed. These three clusters become even more visible when scatter plots of all three measurements per participant are analyzed, as can be seen in the supplement Figure S2:

Figure 1b also shows the separation of three groups based on the following *av.F.min.rel* thresholds derived from the three clusters observed in this scatter plot:

- Slow Walkers: *av.F.min.rel* >0.45**Fg* (blue boxes)
- Fast Walkers: 0.05 *Fg* >*av.F.min.rel* > 0.45**Fg* (purple circles)
- Runners: *av.F.min.rel* < 0.05**Fg* (red dots)

Fg being the body weight (the ground reaction force of the body mass due to earth's gravity, 1g = 9.81 m/s^2).

When grouping *av.v.hor* values according to these three groups Figure 1c shows that each of the three subgroups shows values widely independent of age. The apparent age dependency in the overall group can be explained by a dominant proportion of runners in the age group between 9 and 12 compared to the other age groups where the proportion was more or less equally distributed to all three groups (Figure S2b highlights this effect by displaying the 5th order polynomic interpolation of the average minimum relative force (*av.F.min.rel*) vs. age for both genders).

Therefore, the observed group difference is influenced by the compliance of the different age groups to the given command which aim at maximum performance (maximum average power and maximum horizontal speed) while at the same time eliminate running. Since there is a continuous translation between the movement strategy of "slow walking", "fast walking" and (almost) "running" the used movement strategy relies on the individual interpretation and compliance to the given command. At the same, according to Figures 1 and 2 as well as the supplemental Figure S2 there is a significant impact of the selected movement strategy on the outcome parameters where up to 60% (4SC-Up) and up to 35% (4SC-Dn) of the variations are explained by *av.F.min.rel*. However, performance parameters and not compliance are the aimed outcome parameters of this assessment. To eliminate this additional influence, an obvious alternative approach would have been to allow running, because in this case the selection of the movement strategy would have been in line with increase of performance. However, this would require a staircase of more than 4 steps to allow acceleration and deceleration on the staircase. More steps would limit practical application though, for example due to the required ceiling height for such a staircase and also due to resulting safety issues due to its height. In addition, when applying the test in a cohort of patients running might in many cases not be possible.

For the practical reasons mentioned above, instead we propose to compensate this effect by using the following novelle simple normalization algorithm with the aim to minimize influences due to the selection of the movement strategy while still allowing performance comparison by normalizing to an average "walking" equivalent performance equivalent to the average value of *av.F.min.rel* = 0.5**Fg*.

Therefore, for each parameter a 3rd to 5th order polynomial interpolation was calculated (Figures 1b/2b, dashed line). For each data point the interpolated value was then subtracted and the interpolated value for *av.F.min.rel* = 0.5**Fg* was added. Figures 1c/2c show the resulting data corrected by the proposed algorithm to *av.F.min.rel* = 0.5**Fg*.

The effectiveness of the proposed approach can be seen when comparing Figures 1d) and 1c): Figure 1d shows for 4SC-Up the linear interpolation of *av.v.hor* per group. While the mean value depends on the specific group (from bottom to top: blue line, slow walkers; purple line, fast walkers; red line, runners) there is no obvious age dependency. After the proposed normalization (Figure 1e) the linear interpolation of all three groups are almost identical.

Figure S1. Photographs of the Leonardo Mechanograph® Stair A (left) and Leonardo Mechanograph® GRFP LT device with adjustable bench (right)

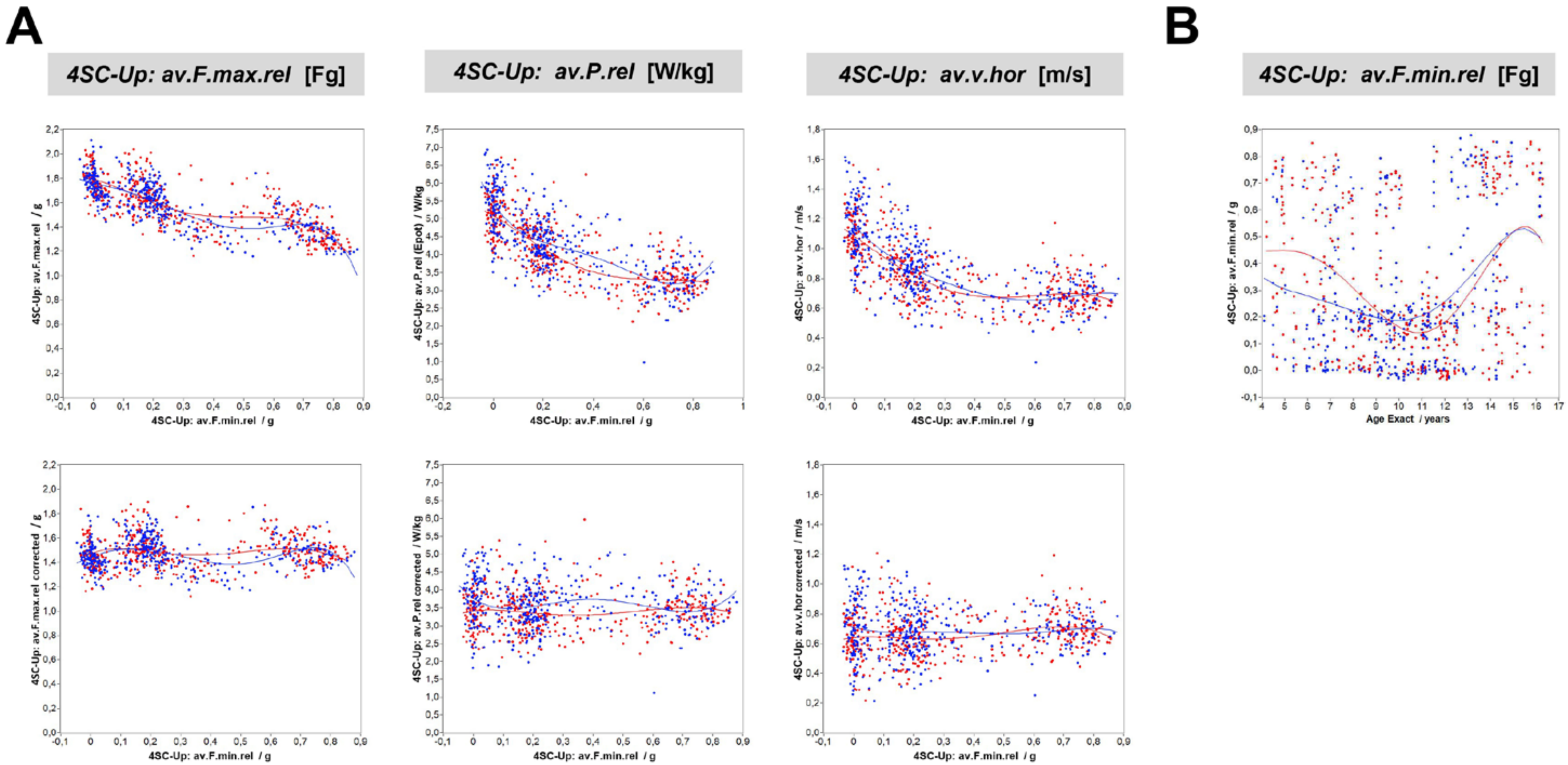

Figure S2. A. Scatterplots of the minimum force during ascent of four steps (av.F.min.rel) and selected mechanographical parameters separated by gender (blue: male; red: female) for all measurements (3 per individual). All parameters show a clear separation of three clusters: runners (*av.F.min.rel*<0.05**Fg*), fast walkers (0.05**Fg* >*av.F.min.rel* >0,45**Fg*), slow walkers (*av.F.min.rel* >0,45**Fg*). B. Visualisation of *av.F.min.rel* results for all age categories. Participants aged 9 to 12 years show lower mean minimal forces, thus containing more, Runners'.

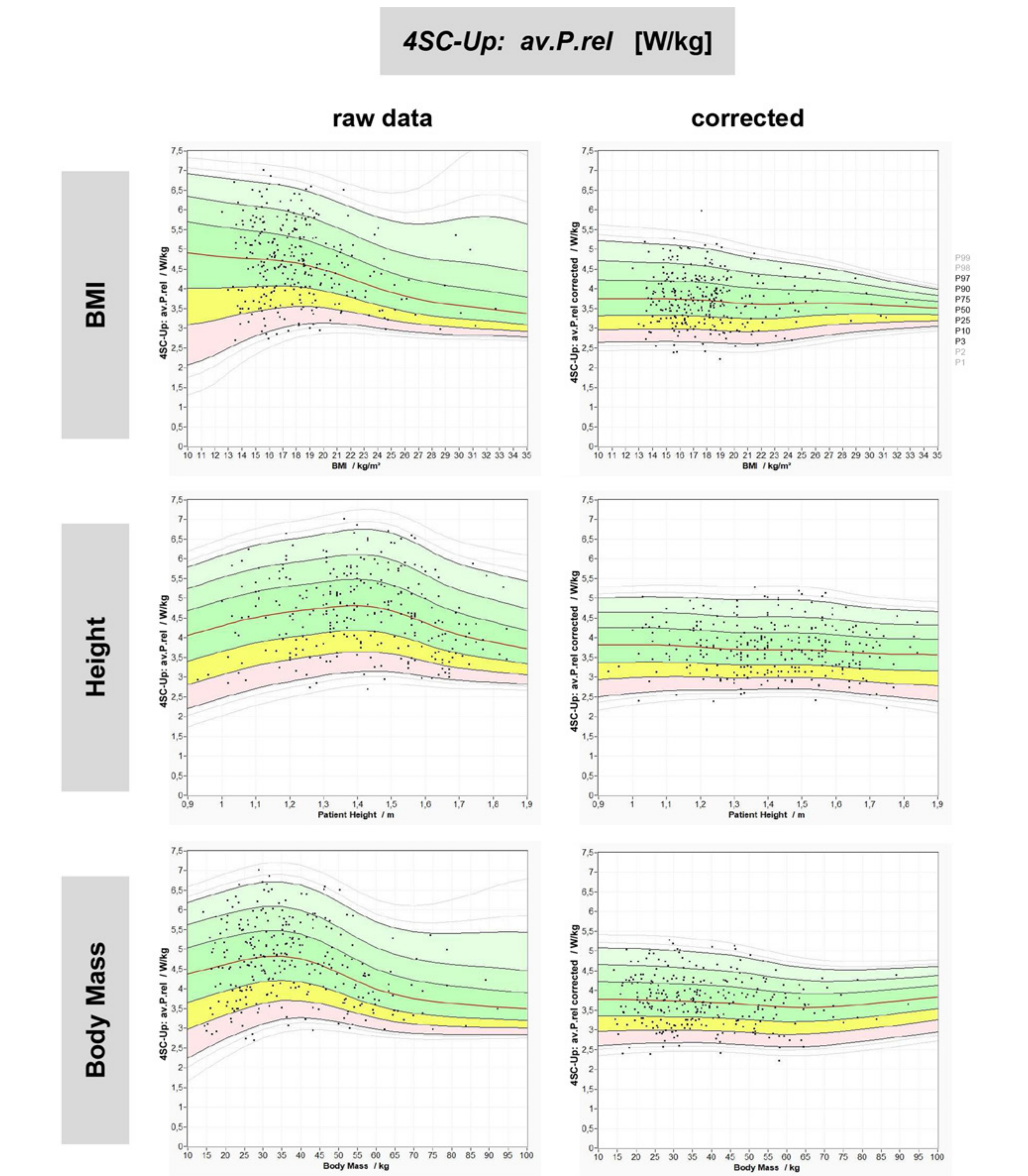

Figure S3. Percentiles of the average power during the ascent of 4 stair climb (*av.P.rel*) according to body mass index (BMI, above), height (middle) and body mass (below) as raw data and corrected for *av.F.min.rel* = 0.5**Fg*.

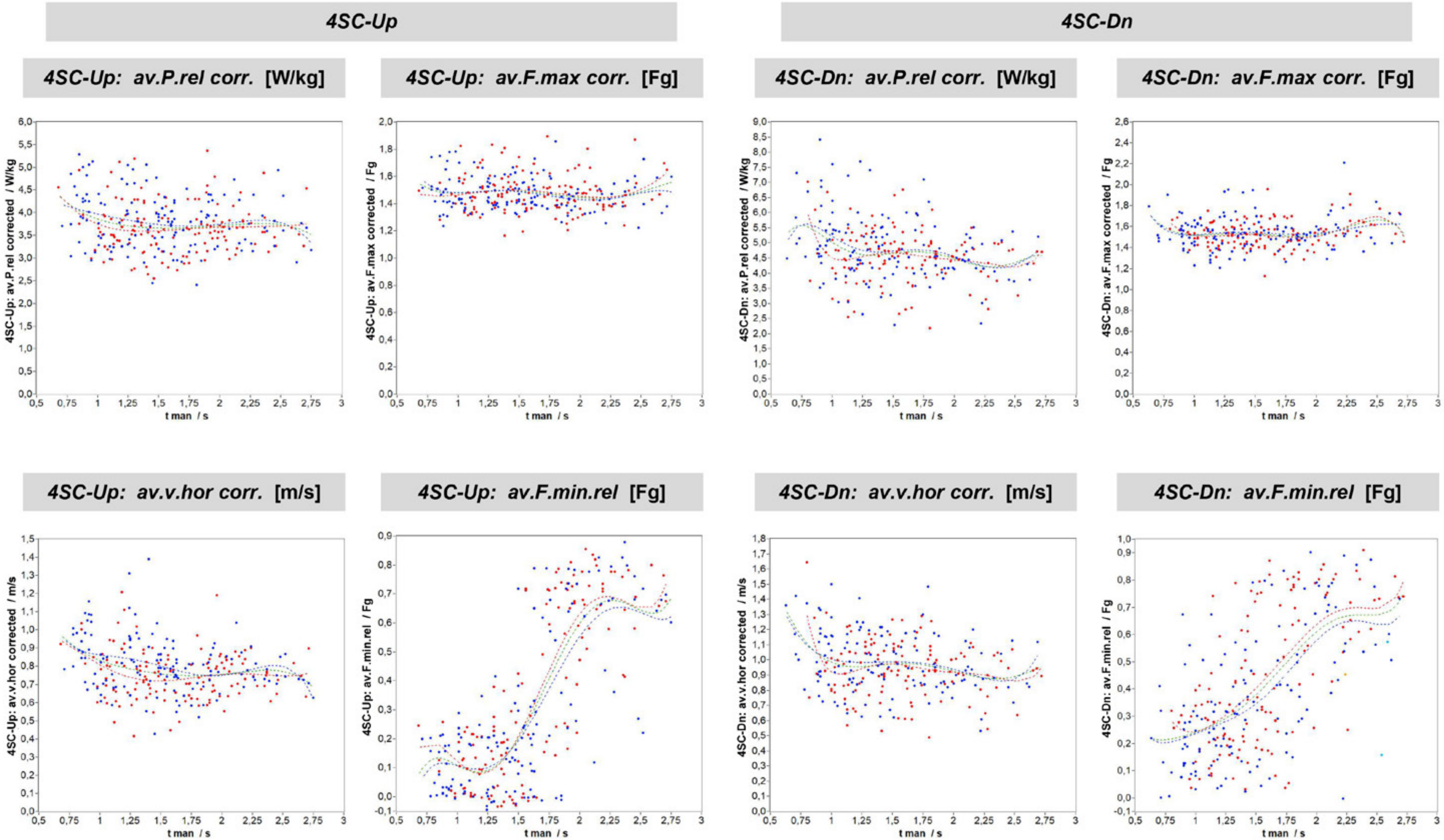

Figure S4. Scatterplot diagrams displaying results of the selected mechanographic parameters during the ascent (left) and descend (right) of 4 stairs and the corresponding manually measured time required for the maneurvre (t.man). Red dots represent female participants and blue dots males. Same applies for interpolation lines. The green interpolation line represents all measurements.

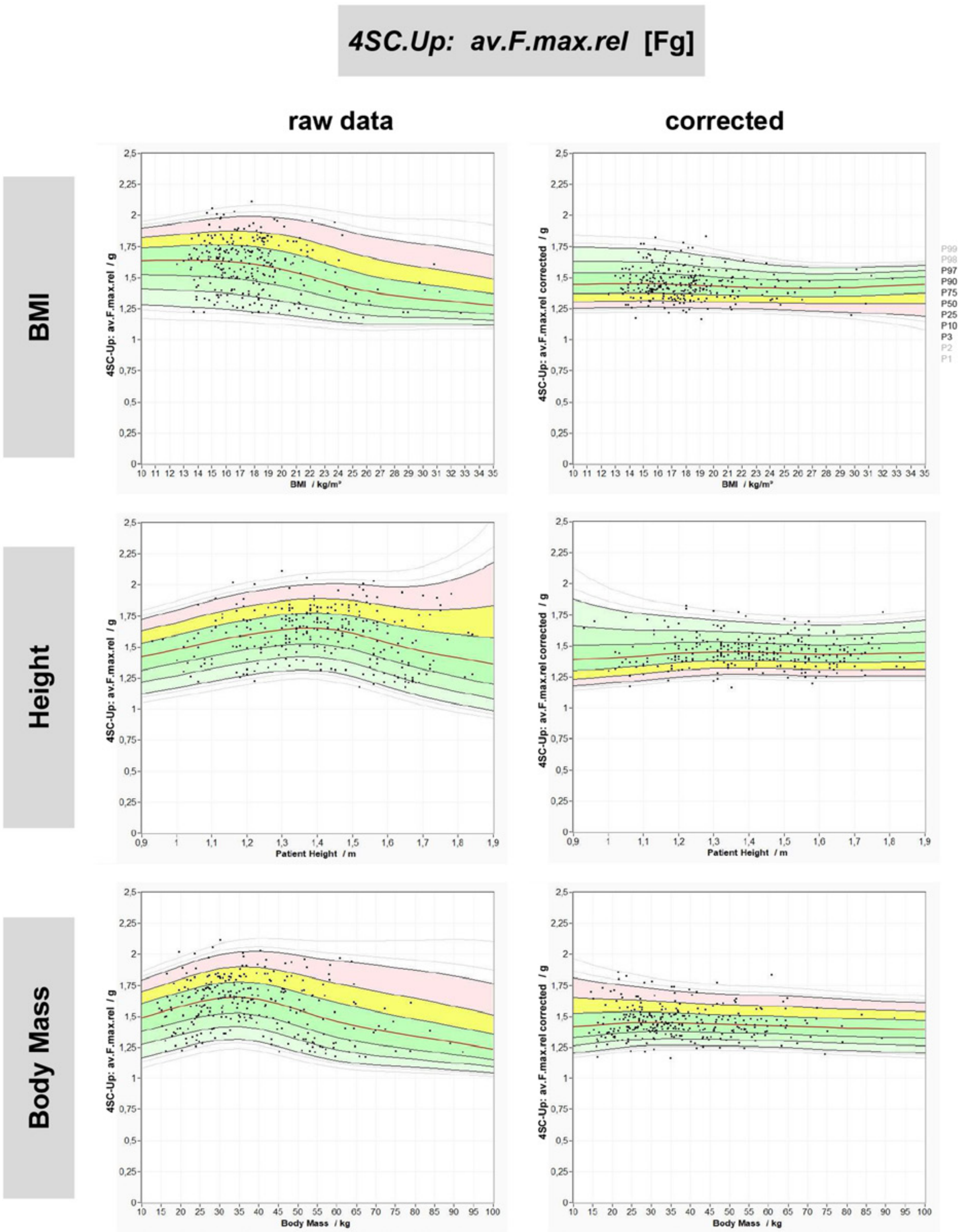

Figure S5. Percentiles of the average maximum force during the ascent of 4 stairs (*Av.F.max.rel*) according to body mass index (BMI, above), height (middle) and body mass (below) as raw data and corrected for *av.F.min.rel* = 0.5**Fg*.

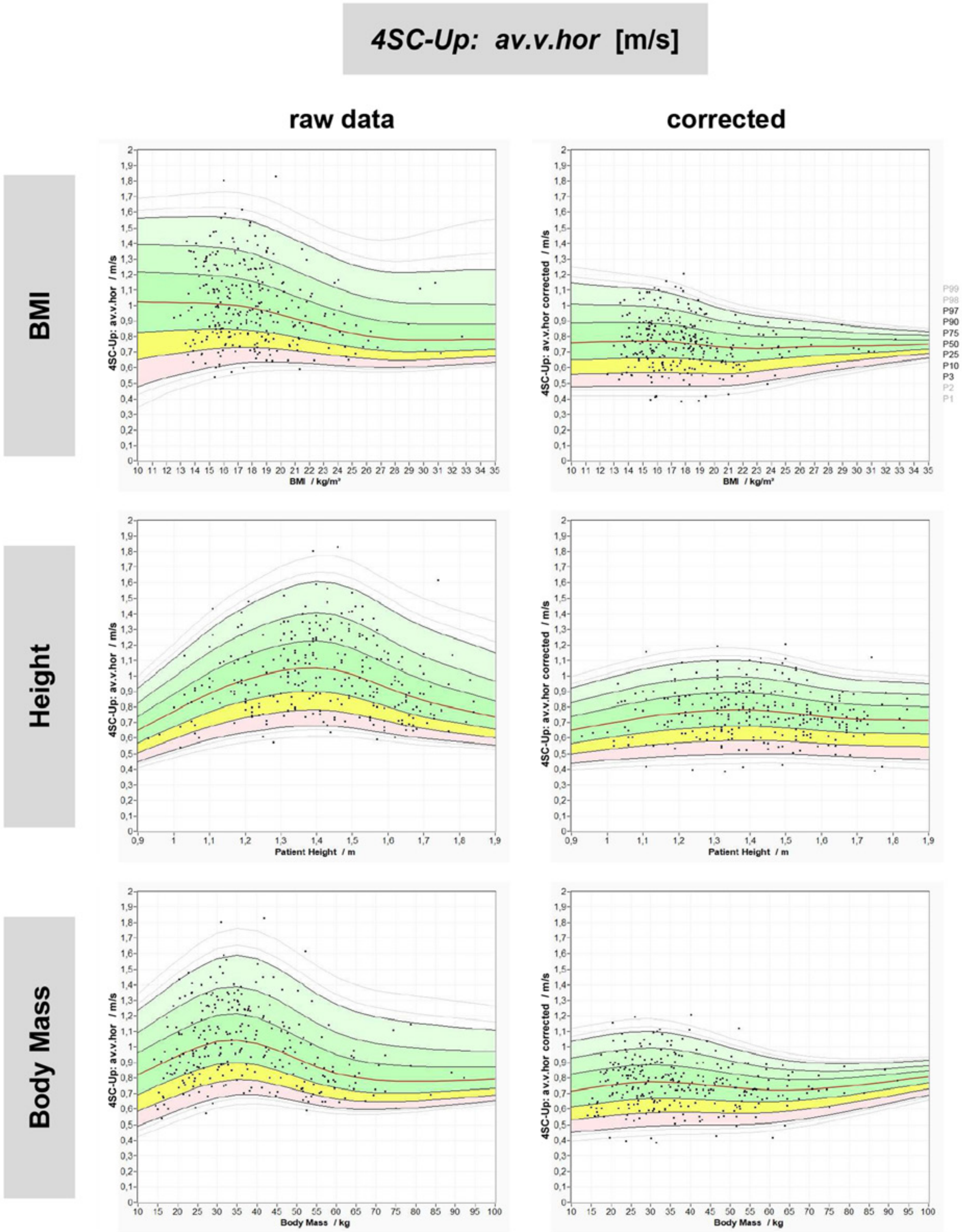

Figure S6. Percentiles of the average maximum horizonta velocity during the ascent of 4 stairs (*Av.V.max*) according to body mass index (BMI, above), height (middle) and body mass (below) as raw data and corrected for *av.F.min.rel* = 0.5**Fg*.

Table S1. Dimensions of available measurement instruments for analysis of the 4 stair climb test.

	Ideal (Sammons Preston) Used in numerous clinical trials	**Standard Leonardo Stair C** Mechanography device - commercial available (Novotec GmbH)	**Custom build stair Leonardo Stair C SP** Mechanography device - used in this study
Height of steps	6` = 15.2 cm	4.7‘ = 12.0 cm	5.9‘ = 15.0 cm
Depth of steps	11` = 27.9 cm	7.5‘ = 19.0 cm	11.0‘ = 28.0 cm
Width of steps	36` = 91.4 cm	31.9‘ = 81.0 cm	31.9‘ = 81.0 cm
Height of railing	24‘ = 61.0 cm	24.4‘ = 62.0 cm	22.4‘ = 57.0 cm
Depth of plateau (upper plateau)	30‘ = 76.0 cm	24.4‘ = 62.0 cm	22.4‘ = 57.0 cm
Depth of plateau (lower Plateau)	-	15.4‘ = 39.0 cm	15.4‘ = 39.0 cm
Steps	4	4	4

Table S2. Standardized protocol of mechanographic assessments.

1. Stair-Climb-Test (SC-Up and SC-Dn)	
Protocol	The test person stands with both legs on the lower platform of the staircase device. Then he is asked to master the four steps up to the upper plateau as quickly as possible. Use of the attached side railing is permitted here (since the force exerted by the arms is also recorded via the floor sensors). In the following. climbing down the stairs starting from the upper plateau is examined for speed in an analogous manner. A break of up to 60 seconds may be inserted between the two subtests (upstairs / downstairs). *Standardized instructions:* *1. "Stand with both feet on the lower plateau"* *2. "Let both arms hang down"* *3. "Without running. climb all the stairs as quickly and as safe as possible as soon as I you hear the start signal. stand still without turning on the upper plateau"* *Turn around:* *4. "Stand with both feet on the upper plateau"* *5. "Let both arms hang down"* *6. "Without running. go down all the stairs as quickly as possible as soon as I you hear the start signal"* The test is carried out three times. provided that the test person's endurance and strength allow it. After each cycle of ascent and descent a break of 30 seconds is mandatory. During the mechanographic measurement. the time taken to climb the stairs is manually measured by the other examiner using a stopwatch. Of the three measurements. the measurement with the best time result for climbing the stairs is used for further evaluation.
2. Grip-Force-Test (GrF)	
Protocol	While sitting. the test person is asked to bend the elbow to 90 ° and to apply maximum force to the measuring device (Leonardo Mechanograph® GF) with the right hand when closing the fist. The measurement is then repeated with the left hand. The measurement is repeated three times. The side dominance (right-handed / left-handed) is queried and documented. *Standardized instructions:* *1. "Sit on the bench. hold your right arm like this (examiner angles elbow to 90 °)"* *2. "Clench your fist as hard as you can as soon as I say GO and hold for 5 seconds"* (Same process for the left hand) A break of 30 seconds is mandatory before the same side is tested again. Of the three measurements. The measurement with the best force result is used for each hand.
3. Chair-Rising-Test (CRT)	
Protocol	Test procedure: The height-adjustable bench on which the test person sits is adapted to the size of the test person and mounted on the mechanograph base plate. The test person is asked to get up as quickly as possible five times in a row from a sitting position (knees bent at 90°) until maximum knee extension is achieved. and then to sit down again as quickly as possible. The arms should be crossed in front of the chest so that the hands touch each other's shoulders. *Standardized instructions:* *1. "Sit on the bench so that the soles of your feet are touching the ground"* *2. "Cross both arms like this (examiner crosses both arms in front of the chest)"* *3. "Get up as quickly as possible until your legs are straight and sit down again as quickly as possible.* Repeat this as quickly as possible five times in a row as soon as you hear the start signal " The test is carried out three times. provided that the test person's endurance and strength allow it. A break of 30 seconds in sitting position is mandatory after each run. Of the three measurements. the measurement with the best time result is used for further evaluation.

Table S3. Results of mechanographic analysis of grip force measurement and chair rising test (CRT). Results refer to the best of 3 performed measurements. Median results and quartiles (P25;P75) are displayed for all parameters.

	All					Male					Female				
	Grip Force		CRT			Grip Force		CRT			Grip Force		CRT		
Age [ys]	n	F.max [N]	n	Av.P.rel [W/s]	Time (manually) [sec]	n	Av.P.max.rel [N]	n	Av.P.rel [W/s]	Time (manually) [sec]	n	Av.P.max.rel [N]	n	Av.P.rel [W/s]	Time (manually) [sec]
4	17	74.0 (57.5;83.0)	16	8.4 (7.5;10.4)	6.4 (5.7;8.7)	6	71.0 (63.5;93.3)	6	8.1 (7.6;12.0)	6.0 (5.4;8.1)	11	75.0 (45.0;80.0)	10	8.5 (7.1;10.1)	6.7 (5.8;9.1)
5	11	98.0 (83.0;108.0)	11	9.72 (9.5;11.6)	5.5 (5.0;6.5)	6	97.0 (82.5;107.3)	6	0.6 (8.9;12.1)	5.3 (4.9;6.7)	5	107.0 (82.5;124.0)	5	10.8 (9.0;11.4)	6.0 (4.4;6.3)
6	25	111.0 (95.5;133.5)	25	11.9 (8.7,13.4)	5.7 (4.3;6.5)	14	114 (98.8;133.3)	14	12.0 (9.3;13.1)	4.9 (4.1;7.0)	11	105.0 (93.0124.0)	11	11.5 (7.5;14.4)	5.7 (4.4;6.3)
7	25	135.0 (120.0;151.0)	25	11.0 (9.6;12.7)	5.7 (5.0;6.5)	13	146 (120.0;172.0)	13	10.2 (8.7;13.2)	5.4 (4.8;7.1)	12	126.0 (118.0;136.5)	12	11.6 (10.6;12.7)	6.0 (5.2;6.2)
8	23	147.0 (137.0;160.0)	23	12.2 (10.4;14.0)	4.7 (4.2;5.6)	12	147.5 (137.3;165.0)	12	13.6 (10.8;16.1)	4.6 (3.8;5.4)	11	147.0 (137.0;160.0)	11	12.0 (10.0;12.9)	4.7 (4.4;5.8)
9	39	142.0 (111.0;171.0)	39	14.2 (12.4;15.1)	5.2 (4.7;6.6)	24	140.0 (112.0;177.5)	24	14.7 (12.8;15.6)	5.2 (4.8;6.3)	15	147.0 (111.0;160.0)	15	13.7 (11.6;14.3)	5.3 (4.4;7.1)
10	25	145.0 (129.5;176.5)	25	14.7 (13.2;16.3)	4.8 (4.2;6.3)	13	166.0 (119.5;191.0)	13	14.7 (13.5;17.1)	4.9 (3.9;6.2)	12	142.5 (128.5;165.5)	12	14.5 (11.8;16.0)	4.8 (4.3;6.4)
11	39	195.0 (171.0;217.0)	39	13.8 (11.9;15.1)	5.2 (4.7;5.9)	21	190.0 (164.5;230.5)	21	14.0 (11.5;15.2)	5.2 (4.4;5.5)	18	199.5 (174.0;217.8)	18	13.5 (12.0;15.2)	5.3 (4.8;6.0)
12	25	216.0 (192.5;228.0)	25	14.0 (12.7;15.5)	5.6 (4.8;6.4)	14	216.5 (211.0;230.8)	14	14.4 (13.2;17.0)	5.3 (4.5;6.2)	11	207.0 (178.0;226.0)	11	13.6 (12.5;14.9)	5.7 (5.1;6.6)
13	19	258.0 (108.0;282.0)	19	13.6 (12.2;15.7)	5.9 (5.4;6.8)	9	258.0 (205.5;277.0)	9	12.3 (10.1;14.6)	5.4 (5.3;7.0)	10	255.5 (229.3;288.0)	10	14.5 (13.1;16.0)	6.0 (5.7;6.7)
14	26	255.5 (230.0;303.8)	26	14.2 (12.5;15.5)	5.7 (4.6;6.5)	9	287.0 (251.0;338.5)	9	14.3 (13.5;16.4)	4.8 (4.4;6.7)	17	242.0 (221.5;294.0)	17	14.0 (11.5;15.1)	5.8 (5.1;6.4)
15	8	274.5 (239.8;325.8)	8	13.3 (12.8;16.0)	5.5 (5.2;6.3)	2	294.0 (242.0;-)	2	15.9 (13.7;-)	5.9 (5.6;-)	6	273.5 (231.8;313.3)	6	12.9 (12.7;14.3)	5.3 (5.1;6.7)
16	6	293.0 (253.0;334.8)	7	14.9 (13.4;15.7)	5.7 (5.2;6.3)	2	336.5 (333.0;-)	3	15.1 (14.8;-)	6.2 (5.3;-)	4	271.5 (247.0;296.0)	4	14.5 (12.2;16.0)	5.7 (5.3;5.9)