Measuring walking impairment in patients with intermittent claudication: Psychometric properties of the Walking Estimated-Limitation Calculated by History (WELCH) questionnaire

Farhad Rezvani [Corresp., 1], **Martin Härter** [1], **Jörg Dirmaier** [1]

[1] Department of Medical Psychology, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Hamburg, Germany

Corresponding Author: Farhad Rezvani
Email address: f.rezvani@uke.de

Objectives. Patient-reported outcome measures can facilitate the assessment of walking impairment in peripheral artery disease patients with intermittent claudication in clinical trials and practice. The aim of this study was to test the psychometric properties of the German version of the 'Walking Estimated-Limitation Calculated by History' (WELCH) questionnaire.

Methods. The assessed properties included feasibility, test-retest reliability, construct validity (i.e., convergent, divergent and known-groups validity) and responsiveness using classic psychometric methods. Psychometric properties were tested as part of a randomized controlled home-based exercise trial for patients with symptomatic peripheral artery disease at Fontaine stage IIA/B.

Results. Analyses were conducted in subgroups of 1696 patients at baseline and 1233 patients at 12-month follow-up (i.e., post-intervention) who completed the WELCH along with a battery of other self-report measures. The WELCH did not exhibit relevant floor or ceiling effects (< 15% achieved lowest or highest possible scores), showed evidence for good test-retest reliability (ICC = 0.81, 95% CI: 0.71 – 0.88) and was found to be well suited for self-completion by patients (< 5% missing data per item). WELCH scores showed moderate to strong correlations with related measures of walking impairment at both time points (Walking Impairment Questionnaire: r = .56 – .74; VascuQoL-25 activity subscale: r = .61–. 66) and distinguished well among patients with poor and high quality of life when adjusting for confounders (t = 13.67, p < .001, d = .96). Adequate divergent validity was indicated by a weaker correlation between the WELCH and general anxiety at both time points (GAD-7: r = -.14 – - .22). The WELCH improved by 6.61 points (SD = 17.04, 95% CI: 5.13-8.10, d = 0.39) in response to exercise treatment and was able to identify large clinically important improvements observed on the walking distance (AUC = .78, 95% CI: .71 – .84) and speed subscales (AUC = .77, 95% CI: .68-.86) of the Walking Impairment Questionnaire.

Conclusions. The WELCH is considered a feasible, reliable and valid patient-reported outcome measure for the measurement of walking impairment in patients with peripheral artery disease. The WELCH showed evidence for responsiveness to changes in walking impairment, yet further studies are warranted to conclusively determine the WELCH's ability to detect intervention effects.

Measuring walking impairment in patients with intermittent claudication: Psychometric properties of the Walking Estimated-Limitation Calculated by History (WELCH) questionnaire

Farhad Rezvani [1], Martin Härter, [1], Jörg Dirmaier [1]

[1] Department of Medical Psychology, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany

Corresponding Author:
Farhad Rezvani, M.Sc.,
University Medical Center Hamburg Eppendorf, Department of Medical Psychology
Martinistrasse 52, 20246 Hamburg, Germany
Email address: f.rezvani@uke.de

Abstract

Objectives. Patient-reported outcome measures can facilitate the assessment of walking impairment in peripheral artery disease patients with intermittent claudication in clinical trials and practice. The aim of this study was to test the psychometric properties of the German version of the 'Walking Estimated-Limitation Calculated by History' (WELCH) questionnaire.

Methods.
The assessed properties included feasibility, test-retest reliability, construct validity (i.e., convergent, divergent and known-groups validity) and responsiveness using classic psychometric methods. Psychometric properties were tested as part of a randomized controlled home-based exercise trial for patients with symptomatic peripheral artery disease at Fontaine stage IIA/B.

Results. Analyses were conducted in subgroups of 1696 patients at baseline and 1233 patients at 12-month follow-up (i.e., post-intervention) who completed the WELCH along with a battery of other self-report measures. The WELCH did not exhibit relevant floor or ceiling effects (< 15% achieved lowest or highest possible scores), showed evidence for good test-retest reliability (ICC = 0.81, 95% CI: 0.71 – 0.88) and was found to be well suited for self-completion by patients (< 5% missing data per item). WELCH scores showed moderate to strong correlations with related measures of walking impairment at both time points (Walking Impairment Questionnaire: r = .56 – .74; VascuQoL-25 activity subscale: r = .61–. 66) and distinguished well among patients with poor and high quality of life when adjusting for confounders (t = 13.67, p < .001, d = .96). Adequate divergent validity was indicated by a weaker correlation between the WELCH and general anxiety at both time points (GAD-7: r = -.14 – - .22). The WELCH improved by 6.61 points (SD = 17.04, 95% CI: 5.13-8.10, d = 0.39) in response to exercise treatment and was able to identify large clinically important improvements observed on the walking distance (AUC = .78, 95% CI: .71 – .84) and speed subscales (AUC = .77, 95% CI: .68-.86) of the Walking Impairment Questionnaire.

Conclusions. The WELCH is considered a feasible, reliable and valid patient-reported outcome measure for the measurement of walking impairment in patients with peripheral artery disease. The WELCH showed evidence for responsiveness to changes in walking impairment, yet further studies are warranted to conclusively determine the WELCH's ability to detect intervention effects.

Introduction

Peripheral Artery Disease (PAD) is a global public health problem affecting an estimated 200 million people worldwide and has become one of the leading causes of disability and death over recent decades. [1-3] The most common symptom is intermittent claudication (IC), which refers to cramping leg pain that is caused by exercise due to insufficient blood flow. [1,4]

Rapid clinical assessment of walking impairment in symptomatic PAD patients is crucial in vascular surgery practice, providing a clinically relevant endpoint from the patient perspective by reflecting walking difficulties in everyday life and therefore considered a better tool for measuring patient-reported walking ability than functional surrogate endpoints. [5] The 'Walking Estimated-Limitation Calculated by History' (WELCH) questionnaire is a brief patient-reported outcome measure (PROM) instrument that requires minimal completion time for assessing walking capacity, with the intention to be used routinely in clinical practice. [6] The WELCH has been translated and cross-validated into various languages [6-10], has shown good feasibility results as it is easy to score and, compared to other PROMs, considered less prone to errors when self-administered by the patient [6,11], while correlating well with treadmill walking. [6,10,12,13]. To be proposed as a routine tool in the future, however, further external validation in larger samples and other languages are required. The purpose of the current study is, therefore, to psychometrically validate the WELCH in a German cohort of symptomatic PAD patients.

Methods

Design

The WELCH was validated as part of a prospective, randomized controlled trial (RCT) evaluating the effectiveness of a 12-month long home-based exercise program for patients with IC, PAD-TeGeCoach. The study protocol was registered and published elsewhere

(ClinicalTrials.gov trial registration: NCT03496948). [14] The study was conducted in accordance with the Declaration of Helsinki and was approved by the ethics committee of the Medical Association of Hamburg (reference number: PV5708). All patients provided written informed consent.

Study population

Approximately 63 000 PAD patients with IC symptoms aged 35–85 with a clinically confirmed ICD-diagnosis of PAD at Fontaine stadium IIa (i.e., IC > 200 meters) or IIb (i.e., IC < 200 meters) within the last 36 months were identified using routinely collected health insurance data from inpatient and outpatient encounters. Patients were excluded, if they had asymptomatic PAD within the last 12 months (Fontaine stadium I) or rest pain within the last 36 months (Fontaine stadium III or IV). As the diagnosis of PAD is often flawed, especially in outpatient settings, participants were interviewed about their IC symptoms prior to enrollment to verify the diagnosis of symptomatic PAD. A sample of 1 982 PAD patients (recruitment rate approx. 3.2%) were enrolled and randomized either into the exercise intervention (PAD-TeGeCoach) or the routine care group (see Figure 1 for RCT flow chart). 11 participants (TeGeCoach n=10; routine care=1) were withdrawn prematurely after randomization (data deletion request n=1, randomized without informed consent n=1, met exclusion criteria n=8, lack of verification of PAD diagnosis n=1), leading to a final sample size of 1971 PAD patients (TeGeCoach n=984; routine care=987).

Measures

Only RCT data from measures relevant to the present study were used and are described in detail elsewhere.[14] While the WELCH was specified in the trial registry as a secondary outcome, the authors failed to include it the study protocol. Notwithstanding this, the (internal and external)

validity of the current study is not expected to be compromised as the trial was registered prospectively (i.e., before recruitment).

Participants received a battery of paper-based questionnaires by mail at each time point and were asked to return them using a prepaid envelope. To maximize return rates, participants who have not returned the questionnaire in time received a postal reminder after 2–4 weeks. All participants were followed up at 12 months, irrespective of whether questionnaires have been returned at baseline. The participants could call the study team when they encountered problems completing the questionnaires.

'Walking Estimated-Limitation Calculated by History' (WELCH) questionnaire

The German version used in this study was requested and made available by the authors of the WELCH, which officially has not yet been psychometrically validated (see supplementary files). The WELCH was forward translated into German by a native-speaking health professional who was not a member of the WELCH development team, and was then closely back translated into French by the authors to ensure appropriate wording. After two rounds of forward and backwards translation, comparing the original and back-translated French versions, a version was reached that was considered acceptable by the authors. The WELCH consists of four items; items 1-3 are eight-point ordinal items, ranging from "impossible" to "3 hours or more", and assess the maximum duration that patients can maintain at different speeds in comparison to friends and relatives (i.e., slower/same/faster). Item 4 is a five-point ordinal item, ranging from "much slower" to "faster", and assesses the usual walking *speed* compared to friends and relatives. The WELCH score is generated by computing the sum of items 1-3, minus one, and multiplying it by the answer of item 4, i.e. [(Item 1 + Item 2 + Item 3) – 1] x Item 4. It is assumed that patients are able to walk at least 30 seconds at low speed so that the sum of the first

three items is never 0. WELCH scores thus range from 0 (i.e., patient is able to walk for a maximum of 30 seconds at slow speed) to 100 (i.e., patient is able to walk 3 hours or more at fast speed). Missing values were handled as indicated by the WELCH authors; for items 1-3 (maximum walking duration), missing values were replaced by the mean of the other two available items (i.e., mean imputation), whereas for item 4, missing values were automatically replaced by 3. [6] The German version of the WELCH can be found in the supplementary files.

Walking Impairment Questionnaire (WIQ)

The Walking Impairment Questionnaire (WIQ) is considered a reliable and valid instrument for assessing walking impairment for different degrees of difficulty across three domains: walking distance, walking speed and stair-climbing. [15-18] WIQ scores are considered responsive to the effect of treatment [15,19], and are strongly correlated with maximum walking distance [5,18], objective measures of walking impairment [16], as well as the ankle-brachial index. [20] The proportion of missing values was <5% for all WIQ items. Item and scale-level missing data were not imputed.

Kings College Vascular Quality of Life questionnaire (VascuQoL-25)

The VascuQoL-25 evaluates PAD-specific quality of life (QOL) and is divided into five subscales: pain, symptoms, activities, social, and emotional. Although designed to measure health-related QOL in PAD patients, all subscale scores and the composite score strongly correlate with functional status outcomes. [21,22] In addition, the activity subscale was suggested to reflect physical functioning in patients with IC according to Wilson and Cleary's model for health-related QOL [23,24], and was therefore considered as a valid comparator instrument. The VascuQoL-25 has also been shown to be responsive to changes in disease severity in PAD

patients. [22,25] The proportion of missing values was < 5% for all VascuQoL-25 items and were mean-imputed.

General Anxiety Disorder Scale (GAD-7)

The GAD-7 was shown to be a reliable and valid self-administered instrument with seven items for screening general anxiety by measuring symptom severity in the last two weeks. [26,27] Items are scored from zero to three on a 4-point scale from 'not at all' to 'nearly every day'. A study conducted in the German general population has found good psychometric properties for the GAD-7. [27] Item-level missing data were not imputed.

Statistical analysis

The psychometric properties of the WELCH were assessed in accordance with the standards set out by the COSMIN group, which guide the development of rigorous methods to investigate psychometric properties of PROMs [28,29]. When available, measurement properties were tested against criteria proposed for good measurement properties of health status questionnaires [30], which were also used in a broad systematic review examining the psychometric validity of PROMs for measuring IC. [24]

Floor and ceiling effects were assessed by examining frequency distributions of each item and were considered to be present if > 15% of the study sample achieved the lowest or highest possible score, respectively. [30] Feasibility of the WELCH was assessed based on the number of missing values in the study sample and per item before imputation. As a rule of thumb, complete case analysis may be used if the proportions of missing data are < 5%, as those missing values

can be considered to be missing at random [31]; therefore, the acceptable proportion of missing data for each item was set at <5%.

The *test-retest* reliability of the WELCH was assessed based on a group of 67 PAD patients at 12-month follow-up who filled it out twice within two weeks under the same condition (i.e., self-completed at home and returned via mail), with patients assumed to be stable during this time period. To determine the level of consistency between these two time points (i.e., stability of repeated measurements), the intraclass correlation coefficient (ICC) and the 95% confidence interval (CI) were calculated using a two-way mixed effects model for single measures with absolute reliability. Test-rest reliability of the WELCH was established when the ICC is > .70. [24,30]

Construct validity of the WELCH was verified with *convergent validity***.** Pearson correlation coefficients, with bootstrapped CIs based on n=1000 samples, were determined between the WELCH and the comparator instruments at baseline and 12-month follow-up (i.e., WIQ, VascuQoL-25 activity subscale), and were deemed satisfactory if there was a strong positive correlation ≥ .50. [24] *Divergent validity* was assessed by testing the association of the WELCH with anxiety, a construct known to be unrelated to symptomatic PAD [32] ; and was considered satisfactory if the correlation between the WELCH and the GAD-7 was weaker than with the comparator instruments at baseline and 12-month follow-up. [24] *Known-groups validity* was determined using a t-test with degrees of freedom adjusted for unequal variances to examine the extent to which the WELCH can significantly discriminate between PAD patients with poor to moderate (VascuQoL-25 score ≤ 4) and high QOL at baseline and 12-month follow-up, given that QOL and walking impairment are known to be associated in patients with IC. [33] The discriminatory ability of the WELCH was also assessed using a multiple linear regression to

control for potential confounders. The model was controlled for all significant confounders reported in Table 1. In the absence of defined quality criteria for health status questionnaires, known-groups validity was also assessed using Cohen's d effect sizes ($0.2 \leq d < .2 \leq$ small, $0.5 \leq d < 0.8$ medium, $d \geq 0.8$ large).

Responsiveness of the WELCH was examined by testing the ability to distinguish patients who have and have not changed after receiving the home-based exercise intervention (TeGeCoach), using the area under the receiver operating characteristics (ROC) curve (AUC) at various threshold settings for minimal clinically important differences (MCIDs) on the WIQ subscales [34]. MCIDs reflect the health status change that patients consider beneficial. [35] For the WIQ subscales, these were previously determined based on an anchor-based method assessing physical function quality of life following a 3-month (home or supervised) exercise intervention, with exercise protocols closely resembling the PAD-TeGeCoach intervention used in this study (i.e., intermittent walking to near maximal claudication pain while using an activity monitor during exercise sessions).[36,37] An AUC of $\geq .70$ was considered to indicate adequate responsiveness. [24,30] In addition, standardized effect sizes (d, baseline – 12-month follow-up) were calculated for the WELCH and the convergent measures.

Analyses were performed using all available data at baseline and during 12-month follow-up. Statistical analyses were performed using SPSS version 25 (IBM Corporation, Armonk, New York, United States). Values of $p < .05$ (two-sided) were considered statistically significant.

Results

Self-reported sociodemographic and clinical information of the PAD patients were collected at both time points and are presented in Table 1, grouped by time point and subgroups to allow to

track the pattern of missing data. Of those enrolled (N= 1971), 1 696 patients returned their questionnaires at baseline (response rate: 86%). 551 patients were lost to 12-month follow-up , while 1145 were followed up through 12-month follow-up. 88 patients returned their questionnaire only at 12-month follow-up, resulting in a sample size of 1 233 patients at 12-month follow-up (response rate: 63%). The response rates fall within the usual range of mail surveys. [38] Reasons for attrition were not identified, but may be attributed to the patient's right to withdraw at any time without having to give a reason and without penalty. The sample sizes, per analysis, using the RCT data are considered excellent for evaluating the psychometric properties of the WELCH.

Feasibility, and floor and ceiling effects

Score distributions and missing values per item before imputation are presented in Table 2. At baseline, the total number of missing values was 44 for item 1 (2.6%), 40 for item 2 (2.4%), 44 for item 3 (2.6%), and 14 for item 4 (0.8%). A total of 1 611 (95.0%) filled out the WELCH completely at baseline, 79 patients (4.7%) filled it out partially, while only 6 WELCH questionnaires were returned completely empty (0.4 %). The number of missing values at 12-month follow-up was similarly low irrespective of study group (Table 2), indicating excellent feasibility. In addition, the number of patients scoring the lowest or highest possible score was < 15% in all items, indicating that there were no floor or ceiling effects irrespective of study group.

Test–retest reliability

A group of 67 PAD patients filled out the WELCH twice within two weeks at 12-month follow-up. The ICC for the WELCH score was 0.81 (95% CI: 0.71 – 0.88), which indicates good test – retest reliability.

Construct validity

Convergent and divergent validity analyses are presented in Table 3. The findings indicate a good convergent validity of the WELCH, as there was a strong positive correlation with the WIQ distance subscale (baseline: n = 1 564, r = 0.65, 95% CI: .62 – .68, p < .001; 12 months: n = 1 110, r = 0.70, 95% CI: .68 – .73, p < .001); a strong positive correlation with the WIQ speed subscale (n = 1 575 r = 0.68, 95% CI: .65 – .71, p < .001; 12 months: n = 1 115, r = 0.72, 95% CI: .69 – .75, p < .001); a strong positive correlation with the WIQ stair climbing subscale (n = 1 590, r = 0.56, 95% CI: .53 – .59, p < .001; 12 months: n = 1 122, r = 0.60, 95% CI: .56 – .63, p < .001); a strong positive correlation with the WIQ total score (n = 1 472, r = 0.70, 95% CI: .68 – .73, p < .001; 12 months: n = 1037, r = 0.74, 95% CI: .72 – .77, p < .001); and a strong positive correlation with the VascuQoL-25 activity scale (n = 1 660, r = 0.61, 95% CI: .58 – .64, p <.001; 12 months: n = 1 193, r = 0.66, 95% CI: .63 – .69, p < .001). Furthermore, there was a weaker correlation between the WELCH and the GAD-7 (baseline: n = 1 629, r = -0.14, 95% CI: -.19 – -.09, p <.001; 12 months: n = 1 162, r = -0.22, 95% CI: -.27 – -.16, p < .001), indicating adequate divergent validity. When separated by study group at 12 months of follow-up (TeGeCoach; routine care), the associations between the WELCH and the comparator instruments were nearly identical (see Table 3), demonstrating satisfactory construct validity regardless of treatment status.

23% of the sample reported poor to moderate QOL at baseline (VascuQoL-25 < 4; n = 383; M = 14.95; SD = 12.20). A significant mean score difference of 17.3 with a very large effect size (d = 0.96; 95% CI: 0.84 – 1.01) was identified between PAD patients with high (n = 1 277; M = 32.23; SD = 19.38) and those with poor to moderate QOL, t (1 001) = 20.91, p <.001, indicating excellent known-groups validity for the WELCH. In multiple linear regression analysis, a

logarithmic transformation was performed to correct for heteroscedasticity. Health-related QOL remained a significant predictor of the WELCH score even after controlling for potential confounders, $t(847) = 13.67$, $p < .001$, which included age, BMI, income, comorbid diseases, medication, gender, education, revascularization and heart rate training. The partial r for predicting WELCH score from health-related quality of life ($pr = 0.43$) was not substantially different from the zero-order Pearson's r without controlling for confounders ($r = 0.47$). Similar results were found at 12-month follow-up irrespective of study group (not reported).

Responsiveness (TeGeCoach home-based exercise)

From baseline to 12-month follow-up, the WELCH [range: 0-100] improved by 6.61 points (SD = 17.04, 95% CI: 5.12-8.10, d = 0.39) in the TeGeCoach group, from 28.86 (SD = 18.98) to 35.47 (SD = 22.29). During the same period, the WIQ distance, speed and stair climbing subscale scores [range: 0-100] improved in the TeGeCoach group by 10.65 (SD = 24.27, 95% CI: 8.44-12.86, d = 0.44), 6.83 (SD = 19.55, 95% CI: 5.05-8.61, d = 0.35) and 5.75 points (SD = 20.51, 95% CI: 3.90-7.60, d = 0.28), respectively. The WIQ total score [range: 0-100] improved by 7.66 points (SD = 17.45, 95% CI: 5.96-9.37, d = 0.44), and the VascuQOL [range: 0-7] by 0.32 points (SE = 0.81, 95% CI: 0.25-0.39, d = 0.40).

Figure 2 presents the ROC curves generated for the WELCH for small (+5% change), moderate (+25% change) and large MCIDs (40% change) on the WIQ, and responsiveness statistics are reported in detail in Table 4. The AUC for small changes was .66 (SE=.02, 95% CI: .62 - .71) for the distance subscale, .64 (SE=.03, 95% CI: .59 - .69) for the speed subscale and .65 (SE=.03, 95% CI: .60 - .70) for the stair climbing subscale. For moderate changes, the AUC was .69 (SE=.03, 95% CI: .64 - .73) for the distance subscale, .64 (SE=.03, 95% CI: .59 - .69) for the speed subscale and .69 (SE=.04, 95% CI: .61 - .77) for the stair climbing subscale. The AUC for

large changes was .78 (SE=.03, 95% CI: .71 - .84) for the distance subscale, .77 (SE=.05, 95% CI: .68 - .86) for the speed subscale and .68 (SE=.05, 95% CI: .58 - .79) for the stair climbing subscale. Another ROC curve has been generated to test the ability of the WELCH to discriminate patients between study groups (routine care vs. TeGeCoach). This showed an AUC of .63 (SE=.02, 95% CI: .59-.66).

Discussion

Further psychometric validation of PROMs that measure walking capacity in PAD patients is required. This study sought to validate the German version of the WELCH, which was developed to address the limitations of existing PROMs for the measurement of walking impairment in PAD patients.

Consistent with previous findings [6], few missing values and >95% completely filled out questionnaires largely support the excellent feasibility of the WELCH in PAD patients. The brief nature of the questionnaire with only four items, and that it can be easily completed without external support, makes it particularly attractive compared to other questionnaires, especially in settings where time pressure is high (e.g. doctor's office). Furthermore, there were no floor or ceiling effects, which enables the WELCH to discriminate equally between symptomatic PAD patients across the entire IC severity spectrum.

In agreement with previous studies [8,11], the WELCH has provided evidence for good psychometric properties in terms of test-retest reliability. This finding is directly related to the usefulness of the WELCH in repeated measurement designs ensuring that scores changes are due to real changes rather than irrelevant artefacts, making an important contribution to its psychometric validity and reliability.

In terms of construct validity, results were in line with previous studies [6,9,10], providing further evidence that the WELCH has good psychometric properties that make it suitable for use in assessing walking impairment living with intermittent claudication. As would be expected, the WELCH demonstrated satisfactory convergent validity, revealing a consistent pattern of moderator to strong correlations with the criterion measures. The associations between the WELCH and others PROMs indicate that the WELCH reflects similar but not identical constructs of walking impairment, with the WELCH showing the highest agreement with measures reflecting walking distance and walking speed (WIQ subscales). Notably, the correlation coefficients between the WELCH score and the WIQ subscales are fairly similar to those observed between the WELCH and treadmill maximum walking distance in previous studies, [6,9,10] which further supports the validity of the WELCH for assessing walking impairment in symptomatic PAD patients. Simultaneously, the WELCH also shows a high correlation with health-related physical functioning (VascuQoL-25 activity subscale), indicating that the WELCH also quantifies the subjective patient experience by reflecting walking limitations in daily life. Furthermore, the WELCH was able to very accurately discriminate between patients with poor and high levels of health-related QOL, demonstrating excellent known-groups validity. Although the relationship between QOL and walking impairment is already well known [33], these results indicate that the WELCH also indirectly reflects aspects of QOL in symptomatic PAD patients, further supporting the WELCH's value as a patient-relevant outcome measure by addressing the impact of PAD on those living with the disease. Finally, as predicted from previous findings [32], the WELCH demonstrated good divergent validity in relation to anxiety symptoms (GAD-7), likewise supporting the validity of the instrument.

318 To date, only few PROMs for PAD patients have been studied in terms their responsiveness,
319 which is a major shortcoming since PROMs are frequently used for measuring the effect of
320 treatment in research and clinical practice, [24] raising doubts on the validity of results. Likewise,
321 evidence on the responsiveness of the WELCH is sparse. [12] In agreement with the construct
322 validity results, the WELCH was found capable to detect large clinically important
323 improvements observed in walking distance and speed following a home-based training regimen,
324 suggesting that the WELCH may be considered responsive to exercise interventions, whereas
325 small to moderate improvements did not generate sufficient change on the WELCH. These
326 findings have direct implications for its use in therapy settings, as they show that the threshold
327 for detecting clinically important effects is relatively high when using the WELCH. The ability
328 to detect a restricted range of clinically meaningful changes (+40%) in response to exercise
329 interventions may limit its utility in clinical settings, particularly since exercise interventions
330 generally have small to moderate, yet clinically meaningful effects on walking impairment. [39,40]
331 The WELCH may therefore be better suited to capture improvements after combined therapies of
332 IC (i.e., exercise therapy plus lower extremity revascularization), as greater improvements in
333 walking performance are usually achieved than after either therapy alone. [41] The limited
334 responsiveness found here is also comparable to the responsiveness of the WELCH after
335 performing revascularization, where a moderate correlation with treadmill maximum walking
336 distance was shown. [12] In addition, the 12-month time interval between the measurements may
337 have also reduced the WELCH's ability to detect improvements after exercise therapy, as the
338 WELCHs responsiveness tends to decrease over time. [12] Despite the promising results reported,
339 the responsiveness of the WELCH in relation to other walking impairment measures (i.e., other
340 PROMs, functional testing), and whether it depends on the time interval between measurements

(i.e., short-term, long-term change), mode of intervention (i.e., invasive, non-invasive) and degree of change (i.e., small, medium, large intervention effects) still remains to be conclusively determined in further studies.

The present study has several strengths compared to previous validation studies, including being the first study to evaluate the psychometric properties of the WELCH in a German clinical population. With a patient to item ratio of 400:1, this is the largest validation study of the WELCH. Furthermore, this is the first study to assess the psychometric properties of a PROM for assessing IC based on the COSMIN checklist, which established guidelines for the psychometric validation of health status PROMs [30]. Although assessing psychometric properties of health status PROMs is common practice, the study quality in the field of PAD is often inadequate [24], which underlines the importance of adopting universal quality criteria. The COSMIN checklist provided a rigorous methodological structure that helped in minimizing methodological bias. It would therefore be useful to use the COSMIN checklist to further evaluate the measurement properties of PROMs for PAD patients.

Several limitations of the study should be noted, including that the translation process did not rigorously comply with the Principles of Good Practice for the Translation and Cultural Adaptation Process for PROMs.[42] Despite this shortcoming, the present study confirms the psychometric validity of the German version of the WELCH, suggesting that the translation and cultural adaptation process can be considered acceptable.

Furthermore, the comparator instruments used to test construct validity are no gold standards, which may have reduced the observed correlations. Notwithstanding this, the WELCH showed high correlations as expected with the comparator instruments, as expected, with the strongest associations on the distance and speed subscales of the WIQ, supporting the validity of the WIQ

in assessing walking impairment [16-19] and thus being well suited as a valid comparator instrument. To provide further evidence for the construct validity of the WELCH, it should also be tested against third-party assessments (e.g. ratings by health professionals) and gold standard measurements (e.g. treadmill testing) in future validation studies.

Conclusions

This article provides evidence that the German version of the WELCH questionnaire is a valid instrument for assessing walking impairment in patients with intermittent claudication. The WELCH, when used appropriately, enables the assessment of walking impairment in PAD patients, while compensating for the existing limitations of existing PROMs. In view of the excellent feasibility and good construct validity, its use can be recommended in clinical settings, as the medical team can quickly gain insight into the PAD patient's walking impairment condition without the need for cumbersome and time-consuming functional assessments. Nonetheless, despite its practicality, the WELCH should be treated with a degree of caution when used to evaluate the benefits of exercise treatments in clinical trials and practice. Alternatively, the WELCH merits consideration in vascular surgery to measure changes evoked by combined treatments (i.e., exercise therapy plus lower extremity revascularization), which, however, remains to be investigated in future studies.

Acknowledgements

We would like to thank the authors of the WELCH questionnaire, who kindly provided us with the German version of the questionnaire. We would also like to thank Finja Mäueler, who assisted us with great enthusiasm and commitment in the preparation of the manuscript.

References

1. Criqui MH, Aboyans V. Epidemiology of peripheral artery disease. *Circ Res.* 2015;116(9):1509-1526.

2. Sampson UK, Fowkes FG, McDermott MM, et al. Global and regional burden of death and disability from peripheral artery disease: 21 world regions, 1990 to 2010. *Glob Heart.* 2014;9(1):145-158 e121.

3. Song P, Rudan D, Zhu Y, et al. Global, regional, and national prevalence and risk factors for peripheral artery disease in 2015: an updated systematic review and analysis. *Lancet Glob Health.* 2019;7(8):e1020-e1030.

4. Aboyans V, Ricco JB, Bartelink MEL, et al. 2017 ESC Guidelines on the Diagnosis and Treatment of Peripheral Arterial Diseases, in collaboration with the European Society for Vascular Surgery (ESVS): Document covering atherosclerotic disease of extracranial carotid and vertebral, mesenteric, renal, upper and lower extremity arteriesEndorsed by: the European Stroke Organization (ESO)The Task Force for the Diagnosis and Treatment of Peripheral Arterial Diseases of the European Society of Cardiology (ESC) and of the European Society for Vascular Surgery (ESVS). *Eur Heart J.* 2018;39(9):763-816.

5. Frans FA, Zagers MB, Jens S, Bipat S, Reekers JA, Koelemay MJ. The relationship of walking distances estimated by the patient, on the corridor and on a treadmill, and the Walking Impairment Questionnaire in intermittent claudication. *J Vasc Surg.* 2013;57(3):720-727 e721.

6. Ouedraogo N, Chanut M, Aubourg M, et al. Development and evaluation of the Walking Estimated-Limitation Calculated by History questionnaire in patients with claudication. *J Vasc Surg.* 2013;58(4):981-988.

7. Abraham P, Godet R, Harbonnier M, Laneelle D, Leftheriotis G, Ouedraogo N. External Validation of the “Walking Estimated Limitation Calculated by History” (WELCH) Questionnaire in Patients with Claudication. *Eur J Vasc Endovasc Surg.* 2014;47(3):319-325.

8. Cucato GG, Correia MdA, Farah BQ, et al. Validation of a Brazilian Portuguese version of the walking estimated-limitation calculated by history (WELCH). *Arq Bras Cardiol.* 2016;106(1):49-55.

9. Abraham P, Godet R, Harbonnier M, Laneelle D, Leftheriotis G, Ouedraogo N. External validation of the "walking estimated limitation calculated by history" (WELCH) questionnaire in patients with claudication. *Eur J Vasc Endovasc Surg.* 2014;47(3):319-325.

10. Tew GA, Nawaz S, Humphreys L, Ouedraogo N, Abraham P. Validation of the English version of the Walking Estimated-Limitation Calculated by History (WELCH) questionnaire in patients with intermittent claudication. *Vasc Med.* 2014;19(1):27-32.

11. Ouedraogo N, Mahe G, Marchand J, Saïdi K, Leftheriotis G, Abraham P. Validation of a new simple questionnaire to “estimate ambulation capacity by history”(EACH) in patients with claudication. *J Vasc Surg.* 2011;54(1):133-138.

12. Henni S, Ammi M, Sempore Y, et al. Treadmill Measured vs. Questionnaire Estimated Changes in Walking Ability in Patients With Peripheral Artery Disease. *Eur J Vasc Endovasc Surg.* 2019;57(5):676-684.

13. Fouasson-Chailloux A, Abraham P, Vielle B, Laporte I, Omarjee L, Ouedraogo N. The correlation of the “Walking Estimated-Limitation Calculated by History”(WELCH)

questionnaire with treadmill maximal walking time is not impaired by age, in patients with claudication. *Qual Life Res.* 2015;24(8):1857-1864.

14. Rezvani F, Heider D, Härter M, et al. Telephone health coaching with exercise monitoring using wearable activity trackers (TeGeCoach) for improving walking impairment in peripheral artery disease: study protocol for a randomised controlled trial and economic evaluation. *BMJ Open.* 2020;10(6):e032146.

15. Regensteiner JG, Steiner JF, Panzer RJ, Hiatt WR. Evaluation of Walking Impairment by Questionnaire in Patients with Peripheral Arterial-Disease. *Clin Res.* 1990;38(2):A515-A515.

16. McDermott MM, Liu K, Guralnik JM, Martin GJ, Criqui MH, Greenland P. Measurement of walking endurance and walking velocity with questionnaire: validation of the walking impairment questionnaire in men and women with peripheral arterial disease. *J Vasc Surg.* 1998;28(6):1072-1081.

17. Sagar S, Brown P, Zelt D, Pickett W, Tranmer J. Further clinical validation of the walking impairment questionnaire for classification of walking performance in patients with peripheral artery disease. *Int J Vasc Med.* 2012;2012.

18. Tew G, Copeland R, Le Faucheur A, Gernigon M, Nawaz S, Abraham P. Feasibility and validity of self-reported walking capacity in patients with intermittent claudication. *J Vasc Surg.* 2013;57(5):1227-1234.

19. Nicolai SP, Kruidenier LM, Rouwet EV, Graffius K, Prins MH, Teijink JA. The walking impairment questionnaire: an effective tool to assess the effect of treatment in patients with intermittent claudication. *J Vasc Surg.* 2009;50(1):89-94.

20. Myers SA, Johanning JM, Stergiou N, Lynch TG, Longo GM, Pipinos, II. Claudication distances and the Walking Impairment Questionnaire best describe the ambulatory limitations in patients with symptomatic peripheral arterial disease. *J Vasc Surg.* 2008;47(3):550-555.

21. Morgan MB, Crayford T, Murrin B, Fraser SC. Developing the Vascular Quality of Life Questionnaire: a new disease-specific quality of life measure for use in lower limb ischemia. *J Vasc Surg.* 2001;33(4):679-687.

22. Mehta T, Venkata Subramaniam A, Chetter I, McCollum P. Assessing the validity and responsiveness of disease-specific quality of life instruments in intermittent claudication. *Eur J Vasc Endovasc Surg.* 2006;31(1):46-52.

23. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA.* 1995;273(1):59-65.

24. Conijn AP, Jens S, Terwee CB, Breek JC, Koelemay MJ. Assessing the quality of available patient reported outcome measures for intermittent claudication: a systematic review using the COSMIN checklist. *Eur J Vasc Endovasc Surg.* 2015;49(3):316-334.

25. de Vries M, Ouwendijk R, Kessels AG, et al. Comparison of generic and disease-specific questionnaires for the assessment of quality of life in patients with peripheral arterial disease. *J Vasc Surg.* 2005;41(2):261-268.

26. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-1097.

27. Lowe B, Decker O, Muller S, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care.* 2008;46(3):266-274.

28. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745.

29. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN study design checklist for Patient-reported outcome measurement instruments. *Department of Epidemiology and Biostatistics.* 2019.

30. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.

31. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res.* 1999;8(1):3-15.

32. Smolderen KG, Hoeks SE, Pedersen SS, van Domburg RT, de L, II, Poldermans D. Lower-leg symptoms in peripheral arterial disease are associated with anxiety, depression, and anhedonia. *Vasc Med.* 2009;14(4):297-304.

33. Golledge J, Leicht AS, Yip L, et al. Relationship Between Disease Specific Quality of Life Measures, Physical Performance, and Activity in People with Intermittent Claudication Caused by Peripheral Artery Disease. *Eur J Vasc Endovasc Surg.* 2020.

34. Gardner AW, Montgomery PS, Wang M. Minimal clinically important differences in treadmill, 6-minute walk, and patient-based outcomes following supervised and home-based exercise in peripheral artery disease. *Vasc Med.* 2018;23(4):349-357.

35. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10(4):407-415.

36. Gardner AW, Parker DE, Montgomery PS, Blevins SM. Step-monitored home exercise improves ambulation, vascular function, and inflammation in symptomatic patients with

peripheral artery disease: a randomized controlled trial. *J Am Heart Assoc.* 2014;3(5):e001107.

37. Gardner AW, Parker DE, Montgomery PS, Scott KJ, Blevins SM. Efficacy of quantified home-based exercise and supervised exercise in patients with intermittent claudication: a randomized controlled trial. *Circulation.* 2011;123(5):491-498.

38. Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol.* 1997;50(10):1129-1136.

39. Golledge J, Singh TP, Alahakoon C, et al. Meta-analysis of clinical trials examining the benefit of structured home exercise in patients with peripheral artery disease. *Br J Surg.* 2019;106(4):319-331.

40. Parmenter BJ, Dieberg G, Smart NA. Exercise training for management of peripheral arterial disease: a systematic review and meta-analysis. *Sports Med.* 2015;45(2):231-244.

41. Biswas M, Capell WH, McDermott MM, et al. Exercise Training and Revascularization in the Management of Symptomatic Peripheral Artery Disease. *Basic to Translational Science.* 2021;6(2):174-188.

42. Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health.* 2005;8(2):94-104.

Figure 1

Flow chart of the study design.

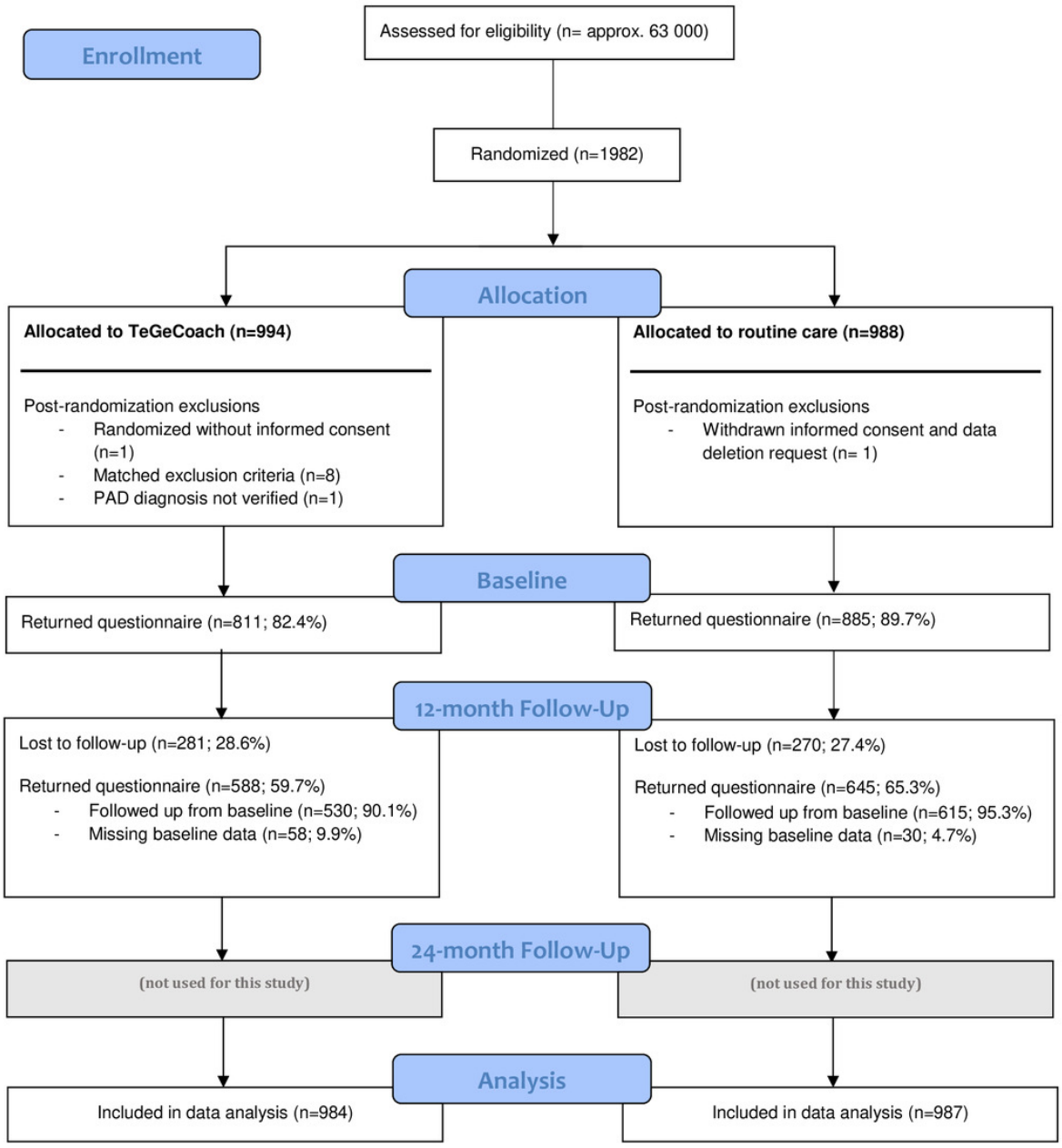

Enrollment
Assessed for eligibility (n= approx. 63 000)
Randomized (n=1982)
Allocation
Allocated to TeGeCoach (n=994)
Post-randomization exclusions
- Randomized without informed consent (n=1)
- Matched exclusion criteria (n=8)
- PAD diagnosis not verified (n=1)
Allocated to routine care (n=988)
Post-randomization exclusions
- Withdrawn informed consent and data deletion request (n= 1)
Baseline
Returned questionnaire (n=811; 82.4%)
Returned questionnaire (n=885; 89.7%)
12-month Follow-Up
Lost to follow-up (n=281; 28.6%)
Returned questionnaire (n=588; 59.7%)
- Followed up from baseline (n=530; 90.1%)
- Missing baseline data (n=58; 9.9%)
Lost to follow-up (n=270; 27.4%)
Returned questionnaire (n=645; 65.3%)
- Followed up from baseline (n=615; 95.3%)
- Missing baseline data (n=30; 4.7%)
24-month Follow-Up
(not used for this study)
(not used for this study)
Analysis
Included in data analysis (n=984)
Included in data analysis (n=987)

Table 1(on next page)

Characteristics of study participants who completed the baseline questionnaire, were lost to 12-month follow-up, were followed through to 12 months, and completed the 12-month follow-up questionnaire.

Sociodemographic characteristics	Baseline			Lost to 12-months follow-up			Followed through 12 months			12-month follow-up*		
	TeGeCoach (n=984)	Routine Care (n=987)	Total (n=1971)	TeGeCoach	Routine Care	Total	TeGeCoach	Routine Care	Total	TeGeCoach	Routine Care	Total
N. of questionnaires received	811	885	1696	281	270	551	530	615	1145	588	645	1233
Sex [a]												
Female	248 (30.6)	281 (31.8)	529 (31.2)	100 (35.6)	94 (34.8)	194 (35.2)	148 (27.9)	187 (30.4)	335 (29.3)	160 (27.2)	199 (30.9)	359 (29.1)
Male	549 (67.7)	597 (67.5)	1146 (67.6)	178 (63.3)	174 (64.4)	352 (63.9)	371 (70.0)	423 (68.8)	794 (69.3)	417 (70.9)	441 (68.4)	858 (69.6)
No information provided	14 (1.7)	7 (0.8)	21 (1.2)	3 (1.1)	2 (0.7)	5 (0.9)	11 (2.1)	5 (0.8)	16 (1.4)	11 (1.9)	5 (0.8)	16 (1.3)
Age (in years) [b]	66.4 (8.6)	66.3 (8.6)	66.3 (8.6)	65.2 (9.4)	63.9 (8.9)	64.6 (9.2)	67.0 (8.2)	67.4 (8.3)	67.2 (8.2)	67.1 (8.3)	67.3 (8.4)	67.2 (8.3)
Minimum - Maximum	35-81	38-81	35-81	38-81	39-81	38-81	35-81	38-81	35-81	35-81	38-81	35-81
BMI [b]	28.1 (5.3)	28.1 (4.8)	28.1 (5.0)	28.4 (5.2)	28.3 (4.9)	28.4 (5.1)	27.9 (5.4)	28.0 (4.7)	27.9 (5.0)	27.8 (5.3)	28.0 (4.7)	27.9 (5.0)
Minimum-Maximum	15.0-75.8	17.0-45.2	15.0-75.8	19.4-54.9	18.3-45.2	18.3-54.9	15.0-75.8	17.0-44.5	15.0-75.8	14.5-75.8	16.8-44.5	14.5-75.8
Education [b] *(multiple choices possible)*												
Apprenticeship	553 (68.2)	613 (69.3)	1166 (68.8)	188 (66.9)	182 (67.4)	370 (67.2)	365 (68.9)	431 (70.1)	796 (69.5)	365 (62.1)	431 (66.8)	796 (64.6)
College	265 (32.7)	297 (33.6)	562 (33.1)	88 (31.3)	73 (27.0)	161 (29.2)	177 (33.4)	224 (36.4)	401 (35.0)	177 (30.1)	224 (34.7)	401 (32.5)
University	140 (17.3)	149 (16.8)	289 (17.0)	35 (12.5)	45 (16.7)	80 (14.5)	105 (19.8)	104 (16.9)	209 (18.3)	105 (17.9)	104 (16.1)	209 (17.0)
Other	68 (8.4)	58 (6.6)	126 (7.4)	30 (10.7)	29 (10.7)	59 (10.7)	38 (7.2)	29 (4.7)	67 (5.9)	38 (6.5)	29 (4.5)	67 (5.4)
No education	44 (5.4)	23 (2.6)	67 (4.0)	22 (7.8)	9 (3.3)	31 (5.6)	22 (4.2)	14 (2.3)	36 (3.1)	22 (3.7)	14 (2.2)	36 (2.9)
Income [b]												
< 500€	16 (2.0)	18 (2.0)	34 (2.0)	11 (3.9)	6 (2.2)	17 (3.1)	5 (0.9)	12 (2.0)	17 (1.5)	5 (0.9)	12 (1.9)	17 (1.4)
500€ to 1000€	63 (7.8)	72 (8.1)	135 (8.0)	30 (10.7)	31 (11.5)	61 (11.1)	33 (6.2)	41 (6.7)	74 (6.5)	33 (5.6)	41 (6.4)	74 (6.0)
1001€ to 1500€	101 (12.5)	114 (12.9)	215 (12.7)	43 (15.3)	45 (16.7)	88 (16.0)	58 (10.9)	69 (11.2)	127 (11.1)	58 (9.9)	69 (10.7)	127 (10.3)
1501€ to 2000€	137 (16.9)	145 (16.4)	282 (16.6)	52 (18.5)	31 (11.5)	83 (15.1)	85 (16.0)	114 (18.5)	199 (17.4)	85 (14.5)	114 (17.7)	199 (16.1)
2001€ to 2500€	153 (18.9)	153 (17.3)	306 (18.0)	45 (16.0)	50 (18.5)	95 (17.2)	108 (20.4)	103 (16.7)	211 (18.4)	108 (18.4)	103 (16.0)	211 (17.1)
2501€ to 3000€	110 (13.6)	132 (14.9)	242 (14.3)	31 (11.0)	37 (13.7)	68 (12.3)	79 (14.9)	95 (15.4)	174 (15.2)	79 (13.4)	95 (14.7)	174 (14.1)
3001€ to 3500€	64 (7.9)	83 (9.4)	147 (8.7)	22 (7.8)	16 (5.9)	38 (6.9)	42 (7.9)	67 (10.9)	109 (9.5)	42 (7.1)	67 (10.4)	109 (8.8)
3501€ and more	99 (12.2)	112 (12.7)	211 (12.4)	26 (9.3)	34 (12.6)	60 (10.9)	73 (13.8)	78 (12.7)	151 (13.2)	73 (12.4)	78 (12.1)	151 (12.2)
No information provided	68 (8.4)	56 (6.3)	124 (7.3)	21 (7.5)	20 (7.4)	41 (7.4)	47 (8.9)	36 (5.9)	83 (7.3)	105 (17.9)	66 (10.2)	171 (13.9)
Marital status [b]												
Single	53 (6.5)	60 (6.8)	113 (6.7)	24 (8.5)	24 (8.9)	48 (8.7)	29 (5.5)	36 (5.9)	65 (5.7)	29 (4.9)	36 (5.6)	65 (5.3)
Married	525 (64.7)	563 (63.6)	1088 (64.2)	159 (56.6)	151 (55.9)	310 (56.3)	366 (69.1)	412 (67.0)	778 (67.9)	366 (62.4)	412 (63.9)	778 (63.1)
Divorced/separated	140 (17.3)	157 (17.7)	297 (17.5)	63 (22.4)	58 (21.5)	121 (22.0)	77 (14.5)	99 (16.1)	176 (15.4)	77 (13.1)	99 (15.3)	176 (14.3)
Widowed	77 (9.5)	90 (10.2)	167 (9.8)	31 (11.0)	29 (10.7)	60 (10.9)	46 (8.7)	61 (9.9)	107 (9.3)	46 (7.8)	61 (9.5)	107 (8.7)
No information provided	16 (2.0)	15 (1.7)	31 (1.8)	4 (1.4)	8 (3.0)	12 (2.2)	12 (2.3)	7 (1.1)	19 (1.7)	70 (11.9)	37 (5.7)	107 (8.7)
Number of children [a] *	1.8 (1.1)	1.6 (1.1)	1.7 (1.1)	1.8 (1.2)	1.5 (1.1)	1.7 (1.1)	1.8 (1.1)	1.7 (1.2)	1.7 (1.1)	1.8 (1.1)	1.7 (1.2)	1.7 (1.1)
Minimum-Maximum	0-7	0-11	0-11	0-7	0-7	0-7	0-6	0-11	0-11	0-6	0-11	0-11
Profession [b] *(multiple choices possible)*												
Employed	227 (28.0)	235 (26.6)	462 (27.2)	84 (29.9)	94 (34.8)	178 (32.2)	143 (27.0)	141 (22.9)	284 (24.8)	143 (24.3)	141 (21.9)	284 (23.0)
Unemployed	33 (4.1)	44 (5.0)	77 (4.5)	18 (6.4)	23 (8.5)	41 (7.4)	15 (2.8)	21 (3.4)	36 (3.1)	15 (2.6)	21 (3.3)	36 (2.9)
Housewife/househusband	21 (2.6)	40 (4.5)	61 (3.6)	7 (2.5)	8 (3.0)	15 (2.7)	14 (2.6)	32 (5.2)	46 (4.0)	14 (2.4)	32 (5.0)	46 (3.7)
Retired	514 (63.4)	543 (61.4)	1057 (62.3)	163 (58.0)	133 (49.3)	296 (53.7)	351 (66.2)	410 (66.7)	761 (66.5)	351 (59.7)	410 (63.6)	761 (61.7)
Retired early	30 (3.7)	22 (2.5)	52 (3.1)	7 (2.5)	10 (3.7)	17 (3.1)	23 (4.3)	12 (2.0)	35 (3.1)	23 (3.9)	12 (1.9)	35 (2.8)
Permanently incapacitated for work	18 (2.2)	27 (3.1)	45 (2.7)	10 (3.6)	13 (4.8)	23 (4.2)	8 (1.5)	14 (2.3)	22 (1.9)	8 (1.4)	14 (2.2)	22 (1.8)
Diseases [a] *(multiple choices possible)*												
Myocardial infarction	114 (14.1)	103 (11.6)	217 (12.8)	35 (12.5)	26 (9.6)	61 (11.1)	79 (14.9)	77 (12.5)	156 (13.6)	85 (14.5)	82 (12.7)	167 (13.5)
Stroke	75 (9.2)	74 (8.4)	149 (8.8)	24 (8.5)	22 (8.1)	46 (8.3)	51 (9.6)	52 (8.5)	103 (9.0)	53 (9.0)	53 (8.2)	106 (8.6)
Metabolism disorder	455 (56.1)	510 (57.6)	965 (56.9)	141 (50.2)	139 (51.5)	280 (50.8)	314 (59.2)	371 (60.3)	685 (59.8)	344 (58.5)	383 (59.4)	727 (59.0)
Angina pectoris	107 (13.2)	117 (13.2)	224 (13.2)	41 (14.6)	31 (11.5)	72 (13.1)	66 (12.5)	86 (14.0)	152 (13.3)	71 (12.1)	90 (14.0)	161 (13.1)
Lung disease	124 (15.3)	147 (16.6)	271 (16.0)	53 (18.9)	39 (14.4)	92 (16.7)	71 (13.4)	108 (17.6)	179 (15.6)	79 (13.4)	117 (18.1)	196 (15.9)
Heart Failure	128 (15.8)	131 (14.8)	259 (15.3)	48 (17.1)	35 (13.0)	83 (15.1)	80 (15.1)	96 (15.6)	176 (15.4)	87 (14.8)	101 (15.7)	188 (15.2)
Hypertension	569 (70.2)	656 (74.1)	1225 (72.2)	198 (70.5)	192 (71.1)	390 (70.8)	371 (70.0)	464 (75.4)	835 (72.9)	409 (69.6)	482 (74.7)	891 (72.3)
Diabetes	214 (26.4)	223 (25.2)	437 (25.8)	75 (26.7)	51 (18.9)	126 (22.9)	139 (26.2)	172 (28.0)	311 (27.2)	151 (25.7)	182 (28.2)	333 (27.0)

Cancer	78 (9.6)	77 (8.7)	155 (9.1)	27 (9.6)	15 (5.6)	42 (7.6)	51 (9.6)	62 (10.1)	113 (9.9)	55 (9.4)	64 (9.9)	119 (9.7)
Drugs [a] *(multiple choices possible)*												
Antihypertensive agents	596 (73.5)	657 (74.2)	1253 (73.9)	199 (70.8)	189 (70.0)	388 (70.4)	397 (74.9)	468 (76.1)	865 (75.5)	440 (74.8)	489 (75.8)	929 (75.3)
Platelet function inhibitor	654 (80.6)	716 (80.9)	1370 (80.8)	210 (74.4)	199 (73.7)	409 (74.2)	444 (83.8)	517 (84.1)	961 (83.9)	491 (83.5)	540 (83.7)	1031 (83.6)
Statins	470 (58.0)	513 (58.0)	983 (58.0)	148 (52.7)	133 (49.3)	281 (51.0)	322 (60.8)	380 (61.8)	702 (61.3)	356 (60.5)	397 (61.6)	753 (61.1)
Revascularization [a]												
Yes	257 (31.7)	242 (27.3)	499 (29.4)	75 (26.7)	60 (22.2)	135 (24.5)	182 (34.3)	182 (29.6)	364 (31.8)	192 (32.7)	185 (28.7)	377 (30.6)
No	422 (52.0)	512 (57.9)	934 (55.1)	154 (54.8)	172 (63.7)	326 (59.2)	268 (50.6)	340 (55.3)	608 (53.1)	310 (52.7)	364 (56.4)	674 (54.7))
No information provided	132 (16.3)	131 (14.8)	263 (15.5)	52 (18.5)	38 (14.1)	90 (16.3)	80 (15.1)	93 (15.1)	173 (15.1)	86 (14.6))	96 (14.9)	182 (14.8)
Group heart rate training [a]												
Yes	110 (13.6)	111 (12.5)	221 (13.0)	31 (11.0)	23 (8.5)	54 (9.8)	79 (14.9)	88 (14.3)	167 (14.6)	85 (14.5)	90 (14.0)	175 (14.2)
No	684 (84.3)	754 (85.2)	1438 (84.8)	244 (86.8)	242 (89.6)	486 (88.2)	440 (83.0)	512 (83.3)	952 (83.1)	492 (83.7)	538 (83.4)	1030 (83.5)
No Information provided	17 (2.1)	20 (2.3)	37 (2.2)	6 (2.1)	5 (1.9)	11 (2.0)	11 (2.1)	15 (2.4)	26 (2.3)	11 (1.9)	17 (2.6)	28 (2.3)
Nationality [a]												
German	763 (94.1)	833 (94.1)	1596 (94.1)	265 (94.3)	248 (91.9)	513 (93.1)	498 (94.0)	585 (95.1)	1083 (94.6)	498 (84.7)	585 (90.7)	1083 (87.8)
Other	20 (2.5)	21 (2.4)	41 (2.4)	9 (3.2)	10 (3.7)	19 (3.4)	11 (2.1)	11 (1.8)	22 (1.9)	11 (1.9)	11 (1.7)	22 (1.8)
No information provided	28 (3.5)	31 (3.5)	59 (3.5)	7 (2.5)	12 (4.4)	19 (3.4)	21 (4.0)	19 (3.1)	40 (3.5)	79 (13.4)	49 (7.6)	128 (10.4)

[a] Categorical variables: n (%)
[b] Quantitative variables: M (SD)

*Information on *education, income, marital status, number of children, nationality* and *occupation* was collected at baseline only, and was not available from the 88 participants who completed only the 12-month follow-up questionnaire. These patients were included in *no information provided*.

Table 2(on next page)

Score distributions and missing values of WELCH items at baseline and 12-month follow-up.

	Baseline						12-month follow-up					
	TeGeCoach (n=811)		Routine care (n=885)		Total (n=1696)		TeGeCoach (n=588)		Routine care (n=645)		Total (n=1233)	
	n	%	n	%	n	%	n	%	n	%	n	%
Item 1												
Impossible	13	1.6	4	0.5	17	1.0	16	2.7	5	0.8	21	1.7
30 seconds	9	1.1	6	0.7	15	0.9	5	0.9	11	1.7	16	1.3
1 minute	27	3.3	25	2.8	52	3.1	10	1.7	21	3.3	31	2.5
3 minutes	98	12.1	112	12.7	210	12.4	43	7.3	78	12.1	121	9.8
10 minutes	239	29.5	240	27.1	479	28.2	118	20.1	160	24.8	278	22.5
30 minutes	204	25.2	234	26.4	438	25.8	133	22.6	155	24.0	288	23.4
1 hour	139	17.1	163	18.4	302	17.8	168	28.6	134	20.8	302	24.5
3 hours or more	60	7.4	79	8.9	139	8.2	73	12.4	59	9.1	132	10.7
Total	789	97.3	863	97.5	1652	97.4	566	96.3	623	96.6	1189	96.4
NAs	22	2.7	22	2.5	44	2.6	22	3.7	22	3.4	44	3.6
Item 2												
Impossible	22	2.7	17	1.9	39	2.3	11	1.9	14	2.2	25	2.0
30 seconds	19	2.3	19	2.1	38	2.2	9	1.5	20	3.1	29	2.4
1 minute	60	7.4	77	8.7	137	8.1	32	5.4	55	8.5	87	7.1
3 minutes	199	24.5	192	21.7	391	23.1	86	14.6	129	20.0	215	17.4
10 minutes	249	30.7	262	29.6	511	30.1	153	26.0	168	26.0	321	26.0
30 minutes	134	16.5	169	19.1	303	17.9	134	22.8	128	19.8	262	21.2
1 hour	82	10.1	96	10.8	178	10.5	106	18.0	78	12.1	184	14.9
3 hours or more	26	3.2	33	3.7	59	3.5	39	6.6	29	4.5	68	5.5
Total	791	97.5	865	97.7	1656	97.6	570	96.6	621	96.3	1191	96.6
NAs	20	2.5	20	2.3	40	2.4	18	3.1	24	3.7	42	3.4
Item 3												
Impossible	104	12.8	111	12.5	215	12.7	47	8.0	90	14.0	137	11.1
30 seconds	62	7.6	71	8.0	133	7.8	33	5.6	52	8.1	85	6.9
1 minute	145	17.9	143	16.2	288	17.0	72	12.2	90	14.0	162	13.1
3 minutes	216	26.6	217	24.5	433	25.5	125	21.3	144	22.3	269	21.8
10 minutes	149	18.4	194	21.9	343	20.2	139	23.6	131	20.3	270	21.9
30 minutes	78	9.6	79	8.9	157	9.3	86	14.6	70	10.9	156	12.7
1 hour	32	3.9	29	3.3	61	3.6	60	10.2	35	5.4	95	7.7
3 hours or more	5	0.6	17	1.9	22	1.3	6	1.0	7	1.1	13	1.1
Total	791	97.5	861	97.3	1652	97.4	568	96.6	619	96.0	1187	96.3
NAs	20	2.5	24	2.7	44	2.6	20	3.4	26	4.0	46	3.7
Item 4												
Much slower	120	14.8	131	14.8	251	14.8	62	10.5	93	14.4	155	12.6
Moderately slow	320	39.5	362	40.9	682	40.2	182	31.0	245	38.0	427	34.6
A bit slower	231	28.5	250	28.2	481	28.4	186	31.6	179	27.8	365	29.6
At the same speed	107	13.2	112	12.7	219	12.9	125	21.3	90	14.0	215	17.4
Faster	23	2.8	26	2.9	49	2.9	23	3.9	17	2.6	40	3.2
Total	801	98.8	881	99.5	1682	99.2	578	98.3	624	96.7	1202	97.5
NAs	10	1.2	4	0.5	14	0.8	10	1.7	21	3.3	31	2.5

1
2

Table 3(on next page)

Correlation between WELCH scores and other measures of walking impairment at baseline and 12-month follow-up.

1

	Baseline			12-month follow-up								
				TeGeCoach			Routine care			Total		
Correlation with WELCH score	**n**	**r**	**Bootstrapped 95%-CI**	**n**	**r**	**Bootstrapped 95%-CI**	**n**	**r**	**Bootstrapped 95%-CI**	**n**	**r**	**Bootstrapped 95%-CI**
WIQ walking distance	1564	.65	.62 – .68	532	.71	.67 – .74	578	.69	.64 – .73	1110	.70	.68 – .73
WIQ walking speed	1575	.68	.65 – .71	531	.73	.69 – .76	584	.71	.67 – .75	1115	.72	.69 – .75
WIQ stair climbing	1590	.56	.53 – .59	535	.60	.55 – .64	587	.60	.54 – .64	1122	.60	.56 – .63
WIQ total	1472	.70	.68 – .73	493	.75	.72 – .79	544	.73	.69 – .76	1037	.74	.72 – .77
VascuQoL Activity	1660	.61	.58 – .64	571	.67	.63 – .71	622	.64	.60 – .68	1193	.66	.63 – .69
GAD-7	1629	-.14	-.19 – -.09	559	-.22	-.28 – -.14	603	-.21	-.28 – -.13	1162	-.22	-.27 – -.16

Table 4(on next page)

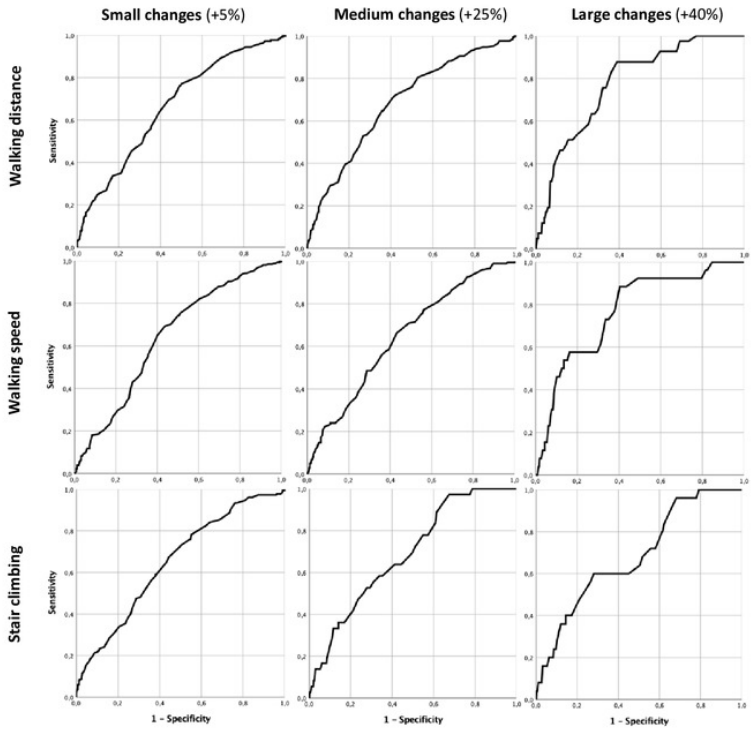
WELCH responsiveness statistics at various thresholds reflecting small, moderate and large changes on the WIQ subscales

	Small (+5% change)			**Moderate** (+25% change)			**Large** (+40% change)		
WIQ subscale	Walking Distance	Walking Speed	Stair climbing	Walking Distance	Walking Speed	Stair climbing	Walking Distance	Walking Speed	Stair climbing
Threshold [a]	≥ 5	≥ 2	≥ 12	≥ 15	≥ 11	≥ 35	≥ 42	≥ 37	≥ 41
AUC	0.66	0.64	0.65	0.69	0.64	0.69	0.78	0.77	0.68
SE	0.02	0.03	0.03	0.03	0.03	0.04	0.03	0.05	0.05
95% CI	0.62-0.71	0.59-0.69	0.60-0.70	0.64-0.73	0.59-0.69	0.61-0.77	0.71-0.84	0.68-0.86	0.58-0.79
n positive	249	256	183	170	175	36	41	26	25
n negative	257	250	323	336	331	470	465	480	481

[a] adopted from Gardner et al., 2018.

Figure 2

WELCH ROC curves for small, moderate and large changes on the WIQ subscales


Small changes (+5%)
Medium changes (+25%)
Large changes (+40%)
Walking distance
Walking speed
Stair climbing
Sensitivity
1 – Specificity