Use of the reliable change index to evaluate the effect of a multicomponent exercise program on physical functions

Haruhiko Sato PT PhD[1*], Masanori Wakida PT PhD [1], Ryo Kubota PT[2,4], Takayuki Kuwabara PT[3,4], Kimihiko Mori PT PhD[1], Tsuyoshi Asai PT PhD[1], Yoshihiro Fukumoto PT PhD[1], Jiro Nakano PT PhD[1], Kimitaka Hase MD PhD [3,4]

[1] Faculty of Rehabilitation, Kansai Medical University, Hirakata, Japan

[2] KMU Daycare Center Kori, Kansai Medical University Kori Hospital, Neyagawa, Japan

[3] Department of Rehabilitation, Kansai Medical University Hospital, Hirakata, Japan

[4] Department of Physical Medicine and Rehabilitation, Kansai Medical University, Hirakata, Japan

Correspondence: Haruhiko Sato, PT, PhD
Faculty of Rehabilitation, Kansai Medical University,
18-89, Uyamahigashimachi, Hirakata, 573-1136, Japan
E-mail: satohar@makino.kmu.ac.jp
Phone: +81-72-856-2251
Fax: +81-72-856-2150
ORCID: 0000-0002-1744-2388

Author contributions: HS contributed to the study conception, design, data analysis, interpretation of data, and preparation of the manuscript. MW contributed to the study conception, material preparation, data collection, and interpretation of data. KM contributed to the study conception and interpretation of data. RK and TK contributed to the study conception, data collection, and interpretation of data. TA, YF, and JN contributed to reviewing and editing. KH contributed to the study conception, reviewing and editing, and project administration. The first draft of the manuscript was written by HS. All authors read and approved the final manuscript.

Running title: Exercise program evaluation using RCI

Abstract

Aims: Using the reliable change index (RCI), we aimed to examine the effect of a multicomponent exercise program on the individual level.

Methods: Overall, 270 adults (mean age, 78 years) completed a multicomponent physical exercise program (strength, aerobic, gait, and balance) for 40 min, 1–2 times per week, continued up to 1 year at a daycare center. Effectiveness was assessed using grip, ankle, knee, and hip strength; Timed Up & Go (TUG); Berg Balance Scale (BBS); gait speed; and 6-min walking distance. These were measured at baseline and every 3 months thereafter. We calculated the RCI using the data between two time points (baseline and at 3, 6, 9, or 12 months) in each participant and then calculated the mean RCI value across the participants. A paired t-test was also employed to evaluate the effect of the intervention as an average-based statistics.

Results: The highest mean RCI values were on ankle plantar-flexion strength, followed by gait speed, hip abduction strength, BBS, knee extensor strength, 6-min walk distance, grip strength, and finally TUG. Paired t-test also revealed significant improvement with moderate effect sizes for ankle plantar-flexion strength (0.504), gait speed (0.413), hip abduction strength (0.374), BBS (0.334), knee extensor strength (0.264), and 6-min walk distance (0.248). A significant but small effect size was seen on TUG (0.183).

Conclusion: The RCI is a convenient method of comparing the effect between different assessments, especially at an individual level. This index can be applied to the use of personal feedback.

Keywords: reliable change index, exercise training, pre–post change, daycare rehabilitation

Introduction

Outpatient daycare rehabilitation is expected to lead to long-term care prevention in older adults.[1] A multicomponent training program including resistance exercises, endurance training, and balance and gait retraining exercises is usually provided as the outpatient daycare rehabilitation from one to three times per week according to the frailty status of the participant [1,2]. Participation in a multicomponent training program has been shown to be effective for older adults [3] in terms of muscle strength [4,5], physical function [4-6], and aerobic capacity [7]. Improvements in muscle strength in conjunction with balance and cardiovascular endurance contribute to improved mobility in older adults [6,8]. In addition, these training programs, which are supervised by rehabilitation specialists such as physiotherapists, are more effective than unsupervised programs such as home exercise [9].

The effects of the exercise intervention, and not only for outpatient daycare rehabilitation, are assessed by comparing pre–post data of physical function tests [6,10]. For multicomponent exercise programs, a variety of physical function tests is used. For example, strength is evaluated as a muscle force measured by a handheld dynamometer, balance function is evaluated as the time needed to perform a balance task (e.g., Timed Up & Go [TUG] test, one-leg standing time), and gait function is evaluated as the gait speed and/or distance walked within a limited period (e.g., 6-minute walking distance).

The pre–post changes in these evaluations cannot be compared directly with each other because of the different distributions and scale units of the data. The amount of change of a certain evaluation test would not indicate the amount of effect itself. The percentage change calculation also has a problem in estimating the effect of an intervention. Even with the same improvement value, the percentage change would be different if the value of the pretest is different. For example, increasing 1 kg for grip strength means a 10% increase for an individual with a baseline of 10 kg, in contrast to a 5% increase for an individual with a baseline of 20 kg [11].

Thus, we focused on the reliable change index (RCI) developed by Jacobson and Truax [12] as an index to determine which physical function component is more effective after participating in a multicomponent exercise intervention, not for a group but rather for each individual. The RCI represents whether the amount of change between pre- and postintervention is statistically large or small; in other words, "Is the change from pretest to posttest meaningful or due to random error? [13]" The RCI is calculated using the individual change and standard error of the measurement for the test. It provides a measure of the change in standardized units, the direction of that change, and whether it is reliable [14]; thus, the RCI is capable of examining the effect of the intervention in each individual.

The effects of the pre–post intervention of the multicomponent exercise program have been examined using the average-based change approach, which evaluates whether a group, as a whole, experiences reliable change [15]. The purpose of this study was to examine the effect of the multicomponent exercise program by the individual-based change approaches such as the RCI, which identifies the specific individual who showed change. We hypothesized that the effect seen from the point of RCI was not different from the effect seen from the point of average-based statistics. We also aim to demonstrate the application of personal feedback to participants using the merit of the RCI.

Methods

Study design and participants

To determine the effects of outpatient daycare rehabilitation on physical function, we retrospectively reviewed records of regular assessments of physical function during outpatient daycare rehabilitation conducted at Kansai Medical University Kori Hospital. This study was approved by the Ethics Committee of the Kansai Medical University (2018251).

Participants were individuals who began attending the daycare center using long-term care insurance between April 2018 and March 2020. The timing of participation varied depending on the participants; however, they continued the multicomponent exercise program for at least 3 months and not more than 1 year. The assessments were performed by experienced physiotherapists (MW, RK, and TK).

Physical function assessment

We used eight different function tests. To assess muscle strength, we used grip strength, strength of hip abduction, knee extension, and ankle plantar-flexion. The Berg Balance Scale (BBS) and the TUG test were used to measure balance function and gait speed, and the 6-minute walking distance was used to evaluate gait function. All tests were assessed at baseline (pretest) and at 3-month intervals after the intervention (a total of 1 year from pretest: 3, 6, 9, and 12 months).

Grip strength: Grip strength was measured using a digital dynamometer (YC III, Tsutsumi, Ltd., Tokyo, Japan). Participants were asked to stand; however, those who were unable to stand were permitted to sit in a relaxed position on a chair without armrests. Measurements were taken once on both the left and right sides while holding the dynamometer with the elbow extended. The maximum values of the left and right sides were used for further analysis.

Hip abduction strength: We measured hip abduction strength using a handheld dynamometer (HHD; µTas F-1, Anima, Tokyo, Japan). With the participant in a supine position with the knee extended, the HHD was positioned 5 cm proximal to the lateral malleolus. The examiner held the HHD, and the participant then exerted maximal force against the HHD [16]. We asked the participant to perform two maximal hip abductor voluntary contractions. The maximal isometric force was used as the representative value for the participant.

Knee extensor strength: Similar to hip abduction strength, we measured knee extensor strength using an HHD. The participant sat on the treatment bed, and the HHD was secured by the belt to the legs of the bed so that the HHD was located at the end of the participant's leg [17]. The participant was asked to perform two maximal knee extensor voluntary contractions. We used the maximal isometric force for further analysis.

Ankle plantar-flexion strength: The participant was in a supine position with the knee extended. The examiner held the HHD and placed it on the plantar surface of the metatarsal head. The participant was then asked to perform two maximum voluntary contractions from a neutral ankle position. The maximum value was used for analysis.

TUG: The TUG test consisted of measuring the time to perform the task of standing up from a chair without hand support, walking 3 m as fast as possible, walking around a cone, returning to the chair, and sitting down again [18]. We used a digital stopwatch to count the time required for participants to complete the task.

BBS: The BBS consisted of 14 items assessing balance tasks such as transfer, unsupported standing, rising from a chair, tandem standing, and one-legged standing. Each task was scored based on a five-point ordinal scale, ranging from 0–4, with 0 indicating the lowest level of function and 4 the highest level of function. The maximum total score of the test is 56 [19].

Gait speed: The participants were asked to walk at the participants' comfortable and safe pace along a 12-m or 6-m walkway. We measured the time at which the participants passed between the 1-m and 11-m marks or between the 1-m and 5-m marks. The participants wore their regular footwear. Two successful trials were recorded, and the faster one was used to calculate the gait speed.

6-minute walking distance: The 6-minute walking distance test is used to measure the distance a person walks as far as possible in 6 min [20]. Two cones were placed 12-m apart, and participants were asked to walk between the cones as long as possible for 6 min.

Multicomponent exercise interventions

A detailed program provided as an outpatient daycare rehabilitation has been published previously [21]. Briefly, the multicomponent exercise program includes lower-limb strengthening exercises (by machine, hip adduction/abduction, leg press), aerobic training (using an ergometer), and task-specific training (gait practice using a single treadmill and gait balance using a dual-belt treadmill [22]) (Supplementary Table 1). Participants were requested to perform a combination of 2 of the 4 types of exercise in one training class. Under the long-term care insurance system, participants attended the training class one to three times a week in the daycare center. The total class time was 40 min and continued for up to 3, 6, 9, or 12 months.

RCI calculation

The RCI is used to evaluate the statistical significance of an individual change between pre- and posttreatment [12,23]. It is based on the standard error from pre- and posttreatment [12,23]. In addition to the original formula proposed by Jacobson, there are several available versions. In the version we used, the equality of the pre- and posttest variances was not assumed [15,23].

$$RCI = \frac{D_i}{\sqrt{\left(Spre\sqrt{1-R_{pre-post}}\right)^2+\left(Spost\sqrt{1-R_{pre-post}}\right)^2}} \quad (1)$$

The D_i describes the individual change in performance of the person undergoing the test. S_{pre} is the standard deviation pretest, and S_{post} is the standard deviation posttest. In addition, $R_{pre\text{-}post}$ is the correlation coefficient between the pre- and posttest.

An RCI value >1.96 indicates reliable improvement at the $p < 0.05$ level, and a value <−1.96 indicates a reliable deterioration. The individual change in performance was calculated based on the measure to obtain the comparative RCI, namely, "pretest minus posttest" in measures aimed at reducing the preintervention score (e.g., TUG), and "posttest minus pretest" in measures aimed at increasing the preintervention score (e.g., grip strength or BBS). We calculated the mean value of the RCI across the participants and the number of participants who showed reliable improvement, reliable deterioration, and no change.

We also calculated %RCI as a readable indicator of reliable change (%RCI = RCI/1.96 × 100). A value greater than 100% indicates reliable improvement, and less than −100% indicates reliable deterioration.

Statistical analysis

For the pre- and postintervention evaluation, we used a paired *t* test for outcomes with normally distributed data and the Wilcoxon rank-sum test for outcomes with nonnormally distributed data. We reported means and standard deviations for normally distributed data and medians and interquartile ranges for nonnormally distributed data. Normality was checked using the Shapiro–Wilk test.

The effect size was used to evaluate the impact of the intervention on the outcome [24]. We used the effect size correlation r_{es}, which can be calculated by the t value of the t test or the z score of the Wilcoxon rank-sum test [25].

$$r_{es} = \sqrt{\frac{t^2}{t^2+df}} \quad (2)$$

$$r_{es} = \frac{Z}{\sqrt{N}} \quad (3)$$

The t describes the t value of the t test, df is the degree of freedom, Z is the z score of the Wilcoxon rank-sum test, and the N is the number of samples. The effect size r_{es} suggested that the $r < 0.2$ would be considered a small effect size, $0.2 \leq r < 0.5$ moderate, and >0.5 large [24].

To compare the RCI value among physical functions, we used the Kruskal–Wallis test followed by the post hoc Dunn multiple comparisons test. The chi-square test was used for comparisons, with the number of people categorized as “improvement,” “deterioration,” or “no change.” We performed these statistical calculations using SPSS version 26 (IBM Corp., Armonk, NY, USA). The significance level was set at 5%.

Results

A total of 293 individuals began using the daycare center between April 2018 and March 2020. Of those, 270 participants (113 men) completed the pretest and posttest. The mean age of participants was 78 years (SD = 8 years; range, 43–95 years). The mean length of time from pretest to posttest was 9 months (SD = 4 months), with 3 months for 60 participants (33 men), 6 months for 44 participants (12 men), 9 months for 44 participants (20 men), and 12 months for 122 participants (48 men).

Table 1 summarizes the changes between pre- and posttest for the eight types of physical function assessment. Significant improvements were observed in most functions except for grip strength. The effect sizes ranged from 0.112–0.504. Ankle plantar-flexion strength had a large effect size (0.504), and gait speed (0.413), hip abduction strength (0.374), BBS (0.334), knee extensor strength (0.264), and 6-minute walk distance (0.248) demonstrated a moderate effect size.

Table 2 summarizes the comparison of the RCI among physical functions and number of people who showed "improved," "deteriorated," and "no change" on each physical function assessment. The highest mean RCI values were for ankle plantar-flexion strength, followed by gait speed, hip abduction strength, BBS, knee extensor strength, 6-minute walk distance, grip strength, and finally the TUG test. The mean RCI of the TUG test was significantly small compared with most other assessments. Similarly, the number of people who were classified as "improved" was greater than the number of people who were classified as "deteriorated" in most assessments except for the TUG. Only two people were classified as "improved," in contrast, five people were classified as "deteriorated" for the TUG test (chi-square test; $p < 0.05$).

Figure 1 shows an example of the results of an individual change between pre- and posttest expressed using %RCI. In this example, the value for the knee extensor strength, BBS, and 6-minute walk distance exceeded +100, which indicates a significant improvement. The value of the TUG test also showed a positive value, but the degree of the value was not as high as the value of the other assessments.

Discussion

We have introduced here an outcome evaluation using the RCI for individuals who participated in a multifunctional exercise program. Most parameters of physical functions showed improvement from both the point of individual-based statistics, such as RCI and average-based statistics. The RCI provides the same information as derived from average-based statistics and allows it to use as personal feedback for the participants. The RCI would be a convenient method to provide functional information for both providers and participants.

Among the eight physical function tests, only the TUG was not so improved from the RCI. The mean value of the RCI across the participants was small (Table 2). This low effect on TUG was also confirmed by the effect size as an average-based statistics, even though the *p*-value was lower than 0.05. A *p*-value itself does not measure the size of an effect or the importance of a result [26]. Thus, both statistical results were not quite different. This close relationship between individual-based statistics and average-based statistics has already been pointed out [15].

Why did the TUG test show less improvement than the other functional assessments? The TUG test is a comprehensive test that assesses the ability to stand up, balance, gait, and agility, which reflects muscle strength. Muscle strength (ankle, knee, and hip), balance (BBS), and gait (gait speed and 6-minute walking distance) showed moderate improvement by the effect size, and significant different effects between the TUG and other assessments were also confirmed by the mean RCI values. Considering this, the TUG should have improved alongside the improvement of strength, balance, and gait. Previous studies have confirmed the smaller effect for the TUG test by the exercise program [5,6,27,28]. This finding might be associated with differences in baseline data and/or the training program itself. The lack of functional activity training such as standing up from a chair and turning direction when walking may have resulted in the discrepancy between the effect of TUG and other assessments. The RCI would have the advantage of evaluating whether a provided exercise program is beneficial by comparing values between the assessments.

We believe that the use of the RCI is convenient for determining whether the change between the pre- and posttest is considered to be large or not so large for each individual. For example, as shown in Figure 1, participants showed positive values of %RCI for all kinds of assessments. However, three assessments exceeded 100, which indicates reliable improvement. An index that makes it easy to understand the magnitude of the change between pretest and posttest would help participants understand their outcome. The %RCI could be applied to use tailored feedback for those who are not specialists.

The RCI has the advantage of being able to show the effects of the intervention for each individual in an easy to understand manner. However, this means a statistically reliable effect, not a clinically meaningful effect. The lower value at the baseline assessment, the larger the margin for improvement;

conversely, the higher the function at the baseline assessment, the smaller the margin for improvement. Thus, participants who demonstrated higher function tended to show a smaller RCI value. If the RCI is used for personal feedback, the rater should take this into consideration to avoid misunderstanding of his or her results.

In conclusion, the RCI, same as effect size, showed a different effect on eight kinds of physical assessment after a multicomponent exercise program. This can be used to evaluate outcomes, not only for a group but also for each individual. The RCI can be used for personal feedback to enable to individual to understand what is improved, is deteriorated, or has no change.

Statements and Declarations

Acknowledgment

We thank the physiotherapists at Kansai Medical University Kori Hospital for their support in running this study. This work was supported by JSPS KAKENHI (grant No. 21K11277).

Data availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Author contributions

HS contributed to the study conception, design, data analysis, interpretation of data, and preparation of the manuscript. MW contributed to the study conception, material preparation, data collection, and interpretation of data. KM contributed to the study conception and interpretation of data. RK and TK contributed to the study conception, data collection, and interpretation of data. TA, YF, and JN contributed to reviewing and editing. KH contributed to the study conception, reviewing and editing, and project administration. The first draft of the manuscript was written by HS. All authors read and approved the final manuscript.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical approval

All procedures in this study that involved human participants were performed following the ethical standards stated in the 1964 Declaration of Helsinki and its later amendments. The study was approved by the Ethical Committee of the Kansai Medical University (No. 2018251).

Informed consent

Written informed consent was obtained from all individual participants included in the study.

References

1. Ministry of Health, Labour and Welfare in Japan (2016) Long-term care insurance of Japan. Available online at: https://www.mhlw.go.jp/english/policy/care-welfare/care-welfare-elderly/dl/ltcisj_e.pdf (accessed March 3, 2022)
2. Hatano E (2012) Long-term Care Insurance in Japan: how physicians are involved in providing rehabilitation for the elderly. Japan Med Assoc J. 55: 231-9.
3. Li Y, Gao Y, Hu S, et al (2022) Effects of multicomponent exercise on the muscle strength, muscle endurance and balance of frail older adults: a meta-analysis of randomised controlled trials. J Clin Nurs doi: 10.1111/jocn.16196
4. Cadore EL, Casas-Herrero A, Zambom-Ferraresi F, et al (2014) Multicomponent exercises including muscle power training enhance muscle mass, power output, and functional outcomes in institutionalized frail nonagenarians. Age (Dordr). 36: 773-85. doi: 10.1007/s11357-013-9586-z
5. Patil R, Uusi-Rasi K, Tokola K, Karinkanta S, Kannus P, Sievanen H (2015) Effects of a multimodal exercise program on physical function, falls, and injuries in older women: a 2-year community-based, randomized controlled trial. J Am Geriatr Soc. 63: 1306-13. doi:

10.1111/jgs.13489

6. Heinrich KM, Crawford DA, Langford CR, Kehler A, Andrews V (2021) High-intensity functional training shows promise for improving physical functioning and activity in community-dwelling older adults: a pilot study. J Geriatr Phys Ther. 44: 9-17. doi: 10.1519/JPT.0000000000000251
7. Tsuji I, Tamagawa A, Nagatomi R, et al (2000) Randomized controlled trial of exercise training for older people (Sendai Silver Center Trial; SSCT): study design and primary outcome. J Epidemiol. 10: 55-64. doi: 10.2188/jea.10.55
8. Webber SC, Porter MM, Menec VH (2010) Mobility in older adults: a comprehensive framework. Gerontologist. 50: 443-50. doi: 10.1093/geront/gnq013
9. Lacroix A, Hortobagyi T, Beurskens R, Granacher U (2017) Effects of supervised vs. unsupervised training programs on balance and muscle strength in older adults: a systematic review and meta-analysis. Sports Med. 47: 2341-61. doi: 10.1007/s40279-017-0747-6
10. Battaglia G, Bellafiore M, Bianco A, Paoli A, Palma A (2010) Effects of a dynamic balance training protocol on podalic support in older women. Pilot Study. Aging Clin Exp Res 22:406-411. https://doi.org/10.3275/6713
11. Dankel SJ, Mouser JG, Mattocks KT, et al (2017) The widespread misuse of effect sizes. J Sci Med Sport. 20: 446-50. doi: 10.1016/j.jsams.2016.10.003
12. Jacobson NS, Truax P (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol. 59: 12-9. doi: 10.1037//0022-006x.59.1.12
13. Ferguson RJ, Robinson AB, Splaine M (2002) Use of the reliable change index to evaluate clinical significance in SF-36 outcomes. Qual Life Res. 11: 509-16. doi: 10.1023/a:1016350431190
14. Zahra D, Hedge C (2010) The Reliable Change Index: why isn’t it more popular in academic psychology? PsyPAG Q. 76: 14-9.

15. Estrada E, Ferrer E, Pardo A (2018) Statistics for evaluating pre-post change: relation between change in the distribution center and change in the individual scores. Front Psychol. 9: 2696. doi: 10.3389/fpsyg.2018.02696
16. Ieiri A, Tushima E, Ishida K, Inoue M, Kanno T, Masuda T (2015) Reliability of measurements of hip abduction strength obtained with a hand-held dynamometer. Physiother Theor Pract. 31: 146-52. doi: 10.3109/09593985.2014.960539
17. Sato H, Kusayanagi K, Kondo Y, Kamide N, Shiba Y, Takashima A (2018) Knee extensor strength assessed using a vertical squat and a simple geometric model to calculate joint torque: an evaluation of validity and clinical utility. Geriatr Gerontol Int. 18: 1125-31. doi: 10.1111/ggi.13299
18. Shumway-Cook A, Brauer S, Woollacott M (2000) Predicting the probability for falls in community-dwelling older adults using the timed up & go test. Phys Ther. 80: 896-903. doi: 10.1093/ptj/80.9.896
19. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL (1992) Clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil. 73: 1073-80.
20. 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: 111-17. doi: 10.1164/ajrccm.166.1.at1102
21. Wakida M, Asai T, Kubota R, et al (2022) Longitudinal effects of physical exercise on health-related outcomes based on frailty status in community-dwelling older adults. Geriatr Gerontol Int. 22: 213-18. doi: 10.1111/ggi.14346
22. Wakida M, Mori K, Kubota R et al (2022) Novel gait training using a dual-belt treadmill in older adults: A randomized controlled trial. Arch Gerontol Geriatr 98:104573.
23. Ferrer R, Pardo A (2014) Clinically meaningful change: false positives in the estimation of individual change. Psychol Assess. 26: 370-83. doi: 10.1037/a0035419

24. Kim HY (2015). Statistical notes for clinical researchers: effect size. Restor Dent Endod. 40: 328-31. doi: 10.5395/rde.2015.40.4.328

25. Yatani K (2014) Statistical methods for HCI research. Available online at: https://yatani.jp/teaching/doku.php?id=hcistats:start#non-parametric_tests (accessed March 30, 2022)

26. Wasserstein RL, Lazar NA (2016) The ASA statement on p-values: context, process, and purpose. Am Stat. 70: 129-33. doi: 10.1080/00031305.2016.1154108

27. Chou CH, Hwang CL, Wu YT (2012) Effect of exercise on physical function, daily living activities, and quality of life in the frail older adults: a meta-analysis. Arch Phys Med Rehabil. 93: 237-44. doi: 10.1016/j.apmr.2011.08.042

28. Raja Adnan RNE, Mat Din H, Ashari A, Minhat HS (2021) Effectiveness of a community-based muscle strengthening exercise program to increase muscle strength among pre-frail older persons in Malaysia: a pilot study. Front Public Health. 9: 610184. doi: 10.3389/fpubh.2021.610184

Table 1: Comparing pretest and posttest measures of physical function.

Physical function	Mean or Median (SD or IQR) Pre-test		Mean or Median (SD or IQR) Post-test		Effect size *r*	*p* value
Grip strength (kg, n = 266)	20.5	(16.1,25.0)	20.5	(17.0,25.5)	0.112	0.068
Hip-abdominal strength (Nm, n = 267)	17.0	(13.5,21.6)	18.9	(15.6,22.8)	0.374	<0.001
Knee-extension strength (Nm, n = 264)	32.3	(26.3,39.2)	34.5	(27.6,40.5)	0.264	<0.001
Ankle plantar-flexion strength (Nm, n = 263)	62.0	(19)	71.0	(20.1)	0.504	<0.001
TUG (seconds, n = 246)	13.1	(10.1,16.9)	12.2	(9.5,16.3)	0.183	0.004
BBS (point, n = 267)	48	(43,52)	50	(44,54)	0.344	<0.001
Gait speed (cm/s, n = 259)	82.3	(25.1)	91.6	(29.5)	0.413	<0.001
6-minute walk distance (m, n = 244)	250	(86)	264	(97)	0.248	<0.001

SD, standard deviation; IQR, interquartile range; TUG, Timed Up & Go test; BBS, Berg Balance Scale.

Table 2: Comparison of reliable change indices and frequency of improved, deteriorated, and no change on physical function measurement

		Grip strength (kg, n = 266)	Hip strength (Nm, n = 267)	Knee strength (Nm, n = 264)	Ankle strength (Nm, n = 263)	TUG (seconds, n = 246)	BBS (point, n = 267)	Gait speed (cm/s, n = 259)	6-minute walking distance (m, n = 244)
RCI	Mean	0.142 [a,b,c]	0.395	0.294	0.584 [d,e]	0.011 [f,g,h,i,j,k]	0.301	0.460	0.260
%RCI	Mean	7.2	20.1	15.0	29.8	0.6	15.3	23.4	13.3
Improved	N	7	15	17	24	5	12	22	11
	%	2.6	5.7	6.4	9.1	2.0	4.5	8.5	4.5
Deteriorated	N	2	2	1	0	2	2	2	2
	%	0.8	0.8	0.4	0.0	0.8	0.7	0.8	0.8
No change	N	257	247	246	239	239	253	235	231
	%	96.6	93.6	93.2	90.9	97.2	94.8	90.7	94.7

RCI, reliable change Index; %RCI, RCI normalized by significant level. Superscripts indicate significant pairwise comparisons in Dann–Bonferroni post hoc analyses.

[a]Grip vs. knee strength.

[b]Grip vs. ankle strength.

[c]Grip vs. gait speed.

[d]Ankle strength vs. knee strength.

[e]Ankle strength vs. 6-minute walking distance.

[f]TUG vs. hip strength.

[g]TUG vs. knee strength.

[h]TUG vs. ankle strength.

[i]TUG vs. BBS.

[j]TUG vs. gait speed.

[k]TUG vs. 6-minute walking distance.

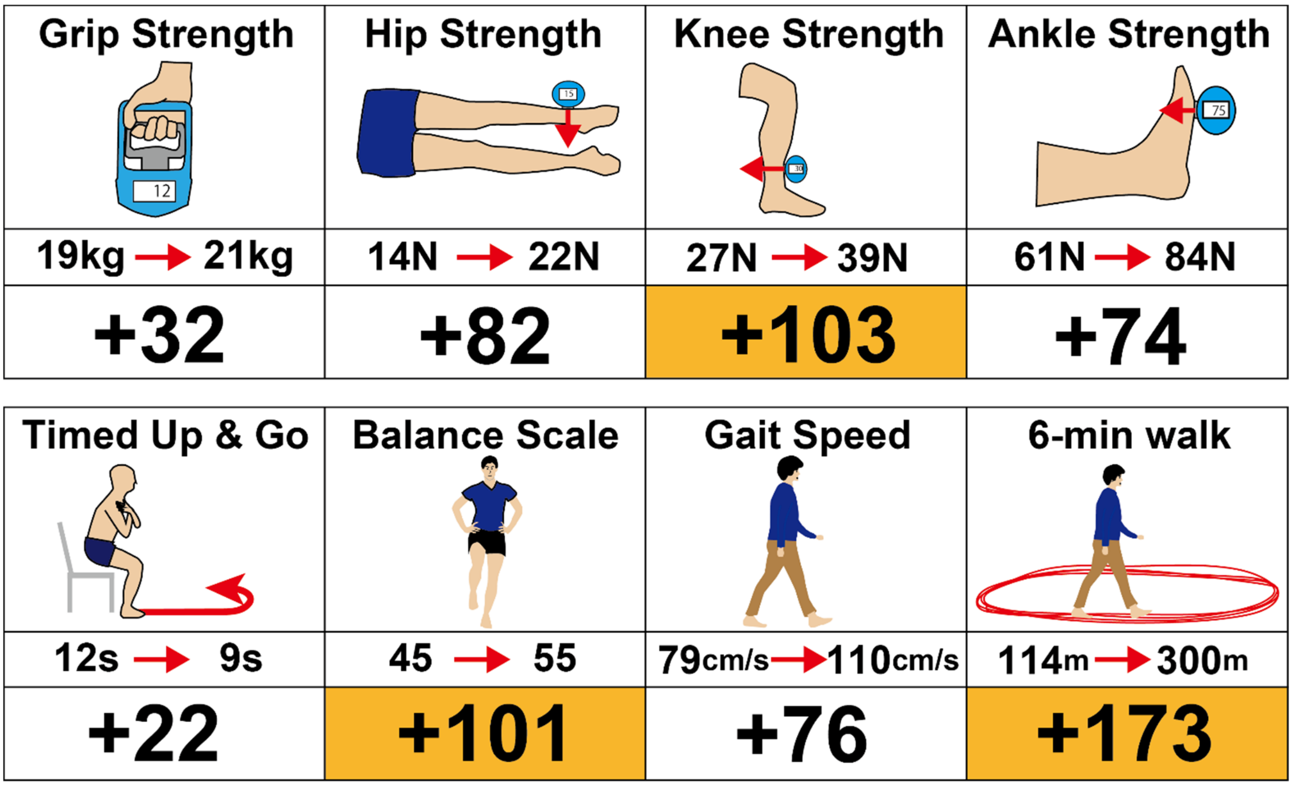

Fig 1. Example of the results of an individual change between pre- and posttest. This image was made as a feedback sheet for a participant. Large-size figures indicate %RCI; the plus indicates "improvement," and the minus sign indicates "deteriorated." A score of >100 means significant improvement