Online supplement

Figure S1. Smartphone application. Left image: the widget on the home screen shows current PA status. Middle image: the bar on the left side combines amount and intensity of steps. The PA goal is met when the vertical stripe (representation of current activity status) is kept in the rising rectangle at all times until the green area is reached. Absolute number of steps and current advice on PA progress are also shown. Right image: the graph shows PA over time.

Figure S2. Website. Anonymized overview of subjects and detailed PA information of a single subject.

A

Power calculations were based on the raw data of a previous study with similar subjects and protocol (1). With effect sizes based on this previous study by Effing et al., analysis with the random intercept, random slope linear mixed 'linear up' model calculations were made in software program PASS 11 for 60, 70, and 80 subjects per group. The power for the time-group interaction for these group sizes was ~72%, ~76%, and ~84% respectively (two-sided test, level of statistical significance: p=0.05). 70-80 subjects per group was deemed sufficient to achieve a satisfactory power.

B

When subjects are repeatedly measured, within subject correlations between the measurement time points are high. In ANOVA these high correlations result in confidence intervals that are too small and to reduced p-values which might subsequently result in a type 1 error. Repeated measures linear mixed modelling (LMM) looks for the accurate correlation structure within the data and then corrects for the tendency of reduced p-values. Additionally, missing data points do not impose any problems in LMM. Therefore, LMM was used (similar to the study of Coxson et al. (2)) to assess the differences between group, the effects of measurement time points, and whether group differences were dependent on measurement time points (=group by time interactions). The chosen approach was a random intercept, random slope method with an unstructured covariance matrix (3).

Average steps/weekday was the primary outcome measure. The primary explanatory parameters were the measurement time points (T0-12) and the group allocation (intervention or usual care group). These two parameters constitute the basic model. Subsequently, other parameters were added to see if they improved the model. These parameters are: age, gender, height, weight, physiotherapy practice, season, temperature, humidity, atmospheric pressure, the number of messages sent, age of the physiotherapist, the number of physiotherapists, long-term care, and all the secondary outcome measures.

The -2log-likelihood (-2LL) is the fit measure of choice in linear mixed modelling: it is a unit-less value and serves the same role as R in linear regression. However, the difference between these two methods is that the value of the -2LL should decrease when the fit between the data and the model improves. The value for the 'extended' model was compared with the basic one and, when a significant decrease was shown based on a χ^2 test, the added parameter was retained. This better fitting model then formed the new 'basic' model to which new parameters were added, and the cycle started again. When the added parameter did not significantly decrease the -2LL value, the parameter was removed from the model and the previous model was maintained as the 'basic' model. This cycle was repeated until the best fitting model was found.

C

Average steps/weekday. The analysis of the primary parameters (time and allocation) showed a significant steps/day decrease over time ($p<0.001$), but no differences between the groups ($p=0.934$). The group by time interaction was also non-significant ($p=0.811$). These data show that both groups decline over time in a similar way. Age and relative humidity could be added to the above basic model as significant parameters. These parameters did not show significant interactions with time or group. These effects are thus constant over time and between groups.

The estimated values map the effect size of a parameter on the dependent variable (see table S1). E.g. an increase in age by one year means that subjects set 94 ± 26 steps/weekday less ($p<0.001$). Compared to

the T12 measurement time point, at T0 subjects take 1178 ± 176 steps/weekday more. By adding all effects we can calculate the estimated steps/weekday for an individual at a chosen time point. When we consider measurement time point T0, a subject who is 60 years old, with 80% humidity, her/his average steps/weekday at that moment would be: 9574 + 1178 + (60 x -94) + (80 x 12) = 6072 steps/weekday. This way we can estimate the number of steps at each of the four measurement time points (see Figure S1).

Table S1. Dependent variable: average steps/weekday. *T12 is the reference value and therefore set at 0.

	estimate	Standard error of the difference	p-value
Intercept	9574		
Time dependent estimate			
T0	1178	176	<0.001
T3	614	176	0.001
T6	472	170	0.006
T12	0*	0*	
Time independent estimate			
Age (years)	-94	26	<0.001
Humidity (%)	12	5	0.027

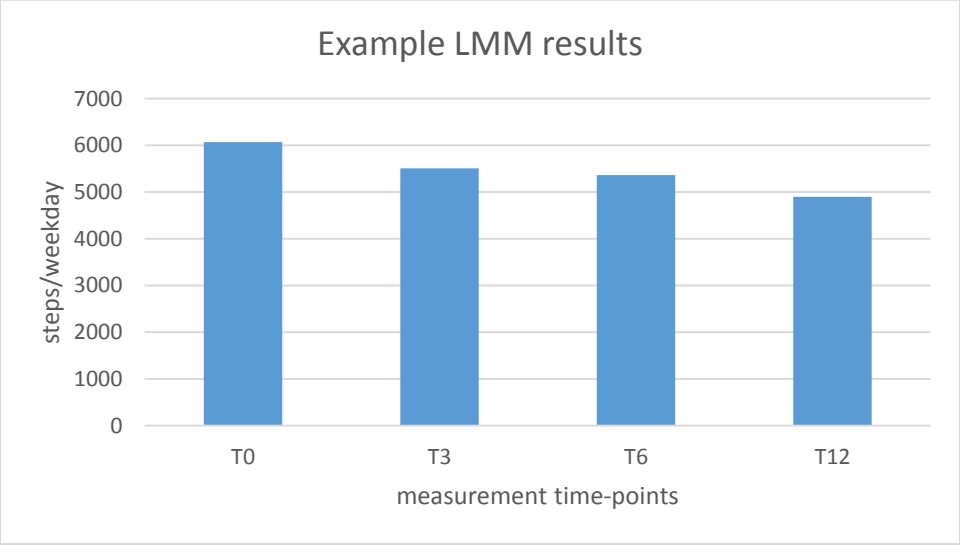


Figure S1. Example of the added effects of the parameters on the dependent variable (steps/weekday) for a subjects of 60 years old, with 80% humidity.

Six-minute walk distance (6MWD). There was no significant decrease in 6MWD over time (p=0.53), and no differences between the groups (p=0.485). The group by time interaction was also non-significant (p=0.585). Gender, BMI, and age could be added to the above basic model as significant parameters. These parameters did not show significant interactions with time or group. These effects are thus constant over time and between groups. The effect sizes of the parameters on the dependent variable are shown in table S2.

Table S2. Dependent variable: 6MWD in meters.

	estimate	Standard error	p-value
Intercept	703.3		
Time independent estimate			
Gender (male)	44.8	12.3	<0.001
BMI (points)	-3.9	1	<0.001
Age (years)	-2.6	0.7	0.001

Average Metabolic Equivalent of Task (Average METs). There was no significant decrease in Average METs over time (p=0.07), and no differences between the groups (p=0.22). The group by time interaction was also non-significant (p=0.36). The 6MWD and BMI could be added to the above basic model as significant parameters. These parameters did not show significant interactions with time or group. These effects are thus constant over time and between groups. The effect sizes of the parameters on the dependent variable are shown in table S3.

Table S3. Dependent variable: Average METs.

	Estimate	Standard error	p-value
Intercept	1.48		
Time independent estimate			
6MWD (meters)	0.0013	0.00028	<0.001
BMI (points)	-0.023	0.0047	<0.001

Lung function

FEV1/FVC. There was a significant decrease of FEV1/FVC over time (p<0.001), but no differences between the groups (p=0.335). The group by time interaction was also non-significant (p=0.908). Secondary parameters did not affect FEV1/FVC significantly. The effect sizes of the various measurement time points on the dependent variable are shown in table S4.

Table S4. Dependent variable: FEV1/FVC in liters. *T12 is the reference value and therefore set at 0.

	estimate	Standard error	p-value
Intercept	0.522		
Time dependent estimate			
T0	0.028	0.005	<0.001
T3	0.021	0.005	<0.001
T6	0.011	0.004	0.006
T12	0*	0*	

FEV1. There was a significant decrease of FEV1 over time (p<0.001), and a significantly higher FEV1 in the intervention group (p=0.05). However, the group by time interaction was non-significant (p=0.508), meaning that there was no effect of the intervention on FEV1. Secondary parameters did not affect the

dependent variable significantly. The effect sizes of the parameters on the dependent variable are shown in table S5.

Table S5. Dependent variable: FEV1 in liters. *T12 is the reference value and therefore set at 0.

	estimate	Standard error	p-value
Intercept	1.53		
Time independent estimate			
Group (intervention)	0.16	0.08	0.05
Time dependent estimate			
T0	0.07	0.02	<0.001
T3	0.04	0.02	0.05
T6	0.02	0.02	0.24
T12	0*	0*	

HRQoL

Dyspnea. The analysis of the primary parameters showed significantly less dyspnea at T3 compared with T12 (p=0.01), but no differences between the groups (p=0.859). The group by time interaction was non-significant (p=0.179). FEV1 could be added to the basic model as significant parameter, but did not show a significant interaction with time or group. The effect sizes of the parameters on the dependent variable are shown in table S6.

Table S6. Dependent variable: Dyspnea in points (1-7; a higher score represents less dyspnea). *T12 is the reference value and therefore set at 0.

	estimate	Standard error	p-value
Intercept	3.9		
Time dependent estimate			
T0	0.13	0.09	0.14
T3	0.22	0.09	0.01
T6	0.11	0.08	0.17
T12	0*	0*	
Time independent estimate			
FEV1 (in liters)	0.49	0.18	0.007

Fatigue. Fatigue did not significantly change over time (p=0.393), and there were no differences between the groups (p=0.879). There was a significant group by time interaction (p=0.018), however, this was probably caused by great variability in the data rather than the intervention (see figure S2). There was no significant difference between the groups at each individual time piont (T0-T12). Dyspnea could be added to the basic model as significant parameter, but did not show a significant interaction with time or group. The effect size of dyspnea on the dependent variable is shown in table S7.

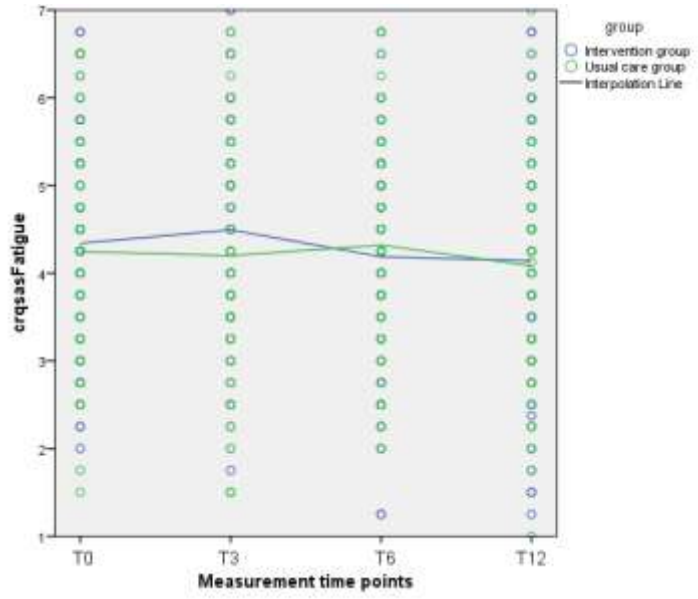


Figure S2. Group by time interaction of fatigue.

Table S7. Dependent variable: Fatigue in points (1-7; a higher score represents less fatigue).

	estimate	Standard error	p-value
Intercept	1.8		
Time independent estimate			
Dyspnea (in points)	0.53	0.04	<0.001

Emotional function. The analysis of the primary parameters showed that emotional function was significantly lower at T0 (p=0.04) and T6 (p=0.02) compared with T12, but there were no differences between the groups (p=0.6). The group by time interaction was not significant (p=0.590). Dyspnea and fatigue could be added to the basic model as significant parameter, but did not show a significant interaction with time or group. The effect sizes of the parameters on the dependent variable are shown in table S8.

Table S8. Dependent variable: Emotional function in points (1-7; a higher score represents a better emotional function). *T12 is the reference value and therefore set at 0.

	estimate	Standard error	p-value
Intercept	2.5		
Time dependent estimate			
T0	-0.12	0.06	0.04
T3	-0.06	0.06	0.29
T6	-0.13	0.06	0.02
T12	0*	0*	
Time independent estimate			
Dyspnea (in points)	0.18	0.03	<0.001
Fatigue (in points)	0.43	0.03	<0.001

Mastery. Mathematical issues prevented us from performing a LMM analysis with this variable as described in the method. To account for this problem we used T0 als covariable instead of dependent variable. Dyspnea and emotional function could be added to the basic model as significant parameter, but did not show a significant interaction with time or group. The effect sizes of the parameters on the dependent variable are shown in table S9.

Table S9. Dependent variable: Mastery in points (1-7; a higher score represents a better mastery). *T12 is the reference value and therefore set at 0.

	estimate	Standard error	p-value
Intercept	1.04		
Time dependent estimate			
T3	0.22	0.08	0.005
T6	0.14	0.08	0.09
T12	0		
Time independent estimate			
Dyspnea (in points)	0.14	0.04	<0.001
Emotional functioning (in points)	0.48	0.05	<0.001

BMI

The analysis of the primary parameters showed that BMI was significantly higher at T6 (p=0.02), but there were no differences between the groups (p=0.223). The group by time interaction was not significant (p=0.458). Height could be added to the basic model as significant parameter, but did not show a significant interaction with time or group. The effect sizes of the parameters on the dependent variable is shown in table S10.

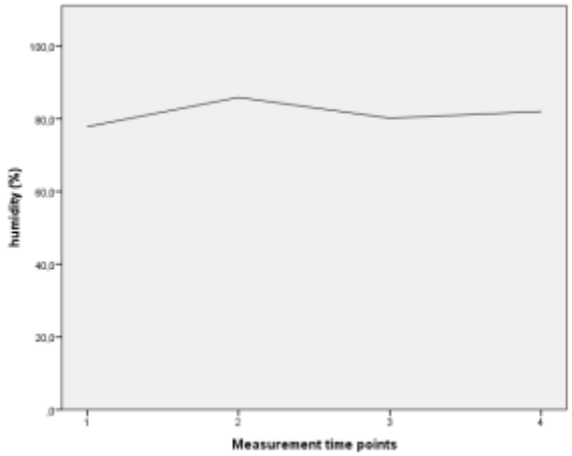
Table S10. Dependent variable: BMI in points. *T12 is the reference value and therefore set at 0.

	estimate	Standard error	p-value
Intercept	52.6		
Time dependent estimate			
T0	-0.09	0.12	0.43
T3	0.03	0.1	0.76
T6	0.18	0.08	0.02
T12	0*	0*	
Time independent estimate			
Height (in meters)	-12.4	6.1	0.045

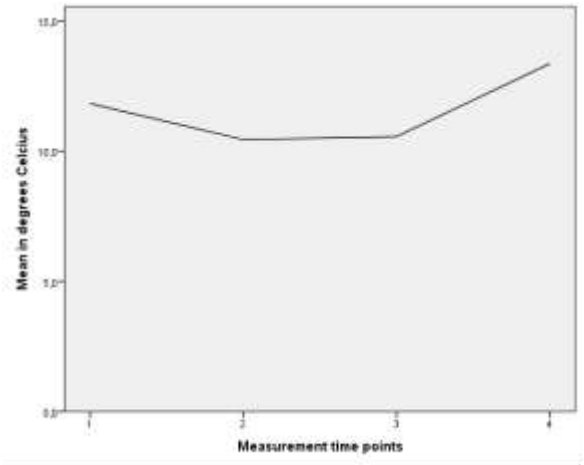
Table S11. Results of the outcome measures.

	T0	T3	T6	T12	Group*time interaction p-values
Average steps/weekday					0.811
I	5824 ± 3418	5285 ± 2669	5267 ± 2669	4819 ± 2883	
U	5717 ± 2870	5318 ± 2889	5139 ± 2804	4950 ± 2634	
I-U	107	-33	128	-131	
6 MWD (in meters)					0.585
I	465 ± 87	475 ± 86	480 ± 86	481 ± 89	
U	459 ± 73	467 ± 70	470 ± 75	471 ± 70	
I-U	6	8	10	10	
Dyspnea (1-7)					0.179
I	4.83 ± 1.25	5.01 ± 1.31	4.88 ± 1.39	4.63 ± 1.49	
U	4.81 ± 1.3	4.8 ± 1.25	4.64 ± 1.33	4.66 ± 1.21	
I-U	0.02	0.21	0.24	-0.03	
Fatigue (1-7)					0.018
I	4.34 ± 1.13	4.49 ± 1.18	4.19 ± 1.31	4.14 ± 1.45	
U	4.25 ± 1.15	4.2 ± 1.29	4.32 ± 1.18	4.08 ± 1.24	
I-U	0.09	0.29	-0.13	0.06	
Emotional function (1-7)					0.590
I	4.95 ± 1.08	5.11 ± 1.06	4.89 ± 1.31	4.94 ± 1.28	
U	4.78 ± 1.24	4.91 ± 1.19	4.87 ± 1.2	4.94 ± 1.17	
I-U	0.17	0.2	0.02	0	
Mastery (1-7)					0.154
I	5.4 ± 1.12	5,49 ± 1.09	5.22 ± 1.27	5.25 ± 1.22	
U	5.32 ± 1.12	5,3 ± 1.1	5.37 ± 1.12	5.12 ± 1.23	
I-U	0.08	0.19	-0.15	0.13	
BMI (kg m^{-2})					0.458
I	27.78 ± 4.86	27,93 ± 4.96	27.93 ± 4.97	27.95 ± 4.96	
U	26.77 ± 5.06	26,69 ± 5.13	27.05 ± 5.07	26.62 ± 5.07	
I-U	1.01	1.24	0.88	1.33	

Data are presented as the mean ± SD. I = intervention group, U = usual care group.



Weather



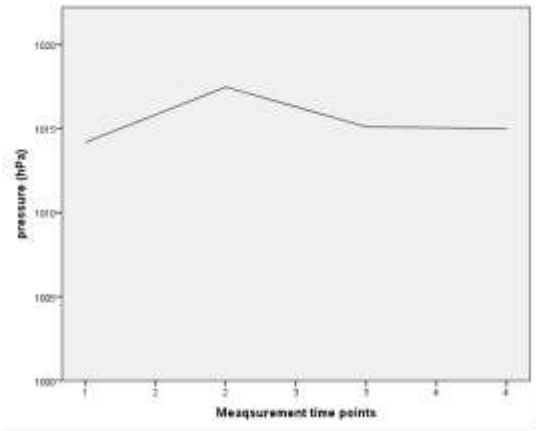


Figure S3: weather conditions at the four measurement time points; humidity, temperature, and atmospheric pressure.

Reference list

(1) Effing T, Zielhuis G, Kerstjens H, van der Valk P, van der Palen J. Community based physiotherapeutic exercise in COPD self-management: a randomised controlled trial. Respir Med 2011 Mar;105(3):418-426.

(2) Coxson HO, Dirksen A, Edwards LD, Yates JC, Agusti A, Bakke P, et al. The presence and progression of emphysema in COPD as determined by CT scanning and biomarker expression: a prospective analysis from the ECLIPSE study. Lancet Respir Med 2013 Apr;1(2):129-136.

(3) Bandyopadhyay S, Ganguli B, Chatterjee A. A review of multivariate longitudinal data analysis. Stat Methods Med Res 2011;20(4):299-330.