



Trajectories of Physical Work Capacity in Early Symptomatic Osteoarthritis of Hip and Knee: Results from the Cohort Hip and Cohort Knee (CHECK) Study

H. J. Bieleman¹ · R. Stewart² · M. F. Reneman² · W. M. van Ittersum³ · C. P. van der Schans³ · K. W. Drossaers-Bakker⁴ · F. G. J. Oosterveld¹

Published online: 27 August 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Purpose To evaluate the 5-year course of physical work capacity of participants with early symptomatic osteoarthritis (OA) of the hip and/or the knee; to identify trajectories and explore the relationship between trajectories and covariates. **Methods** In a prospective cohort study, physical work capacity was measured at baseline, using a test protocol (functional capacity evaluation) consisting of work-related physical activities. Participants were invited to participate in 1, 2 and 5 year follow-up measurements. Multilevel analysis and latent classes analysis were performed, in models with test performances as dependent variables and age, sex, work status, self-reported function (Western Ontario McMasters Arthritis Scale—WOMAC), body mass index (BMI) and time as independent variables. Multiple imputation was used to control for the influence of missing data. **Results** At baseline and after 1, 2 and 5 years there were 96, 64, 61 and 35 participants. Mean (SD) age at baseline was 56 (4.9) years, 84% were females. There was no statistically significant change in test performances (lifting low and high, carrying, static overhead work, repetitive bending, repetitive rotations) between the 4 measurements. Male sex, younger age and better self-reported function were statistically significant ($p < 0.05$) determinants of higher performance on most of the tests; having a paid job, BMI and progression of time were not. Three trajectories were identified: ‘weak giving way’, ‘stable and able’, and ‘strong with decline’. **Discussion** In subgroups of participants with early symptomatic OA, determined by age, sex and self-reported function, physical work capacity seems to be a stable characteristic over 5 years.

Keywords Osteoarthritis · Hip · Knee · Physical work capacity · Cohort · FCE

Introduction

Osteoarthritis (OA) of the hips and the knees can have a severe impact on physical activities in daily life, for example walking and climbing stairs may become limited. Activities in sports [1] and in work may also be affected [2, 3]. Although the onset of OA is usually after the age of 45 and a significant increase in prevalence is seen after 60, a possible impact of OA in people in the working age has been recorded [4]. Among the Canadian OA population, around 50% of non-employment due to illness was associated with OA. Estimations for the size of the working age population with OA increased from 1.5 million in 2010 to 1.7 million in 2031 [5]. In the Netherlands, in the age categories 45–49, 50–54, 55–59 and 60–64 years the prevalence of OA increases from 21.5–41.0–70.4 to 114.3 per 1000 for males and 31.8–70.9–120.8 to 186.8 per 1000 for females [6]. As in Canada, due to demographic changes here also a large increase of the impact of knee and hip OA on

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10926-018-9809-3>) contains supplementary material, which is available to authorized users.

✉ H. J. Bieleman
h.j.bieleman@saxion.nl

¹ Saxion Universities of Applied Sciences, P.O. Box 70.000, 7500 KB Enschede, The Netherlands

² University Medical Centre, University of Groningen, Groningen, The Netherlands

³ Hanze University of Applied Sciences, Groningen, The Netherlands

⁴ Medical Spectrum Twente, Enschede, The Netherlands

work-related sickness is expected [7]. The course of OA and its impact on functioning is generally characterized by slow deterioration in the early phase [8, 9] but some individuals may show different trajectories [10, 11]. Age, body mass, pain, reduced knee flexion, number of comorbidities, lower vitality or avoidance of activities may influence the course of physical functioning in osteoarthritis [9].

The onset of work restrictions in OA is independently associated with severe joint pain and functional limitation, number of affected body sites but also with area employment deprivation [12]. Being male, pain interference with function and lower co-worker support were predictive for work loss in OA, but the extent of arthritis, co-morbidity, obesity or psychological or other job factors were not [13]. Explaining the impact of OA on work in relation to the course of OA in time appears to be a complex, multifactorial matter.

Whether individuals can meet the physical demands of their job basically depends on the physical load in work and the physical work capacity of the individual. Several tools and methods are available for job load analysis [14] and for capacity evaluation [15]. An imbalance between load and capacity may be a risk for the development of limitations and disability. This apparently obvious assumption has not been tested for OA in relation to work. It has been tested for low back and upper extremity disorders, but the assumption was not confirmed [16]. A limited number of cross-sectional studies have been performed that analyze the physical work capacity of groups of people [17–19] but no longitudinal studies applying performance-based measurements are available. Studies with a longitudinal design are therefore necessary to analyze the course of OA and its impact on work capacity, and to identify (modifiable) predictors for this course.

Insight into the course of work capacity of employees with OA is needed, because the rising prevalence of OA with increasing age is expected to amplify the decline in physical functions with age [20–23], although considerable individual variation exists. Some people may demonstrate rapid decline where others may remain stable for years. This insight may generate recommendations for sustainable work participation. The primary objective of this study was to evaluate the 5 years course of functional capacity in people with early OA of the hip and/or the knee. Secondary objectives were to identify factors correlated to this course and (referring to individual differences) to identify trajectories (latent classes) of participants with different courses.

Methods

Design

The study was performed in a sample of 157 out of 1003 participants from the prospective Cohort Hip and Cohort

Knee study [24]. It concerned an additional study in a convenience sample consisting of the participants of 2 of the 10 CHECK centers, Groningen and Enschede, in the context of the collaboration between these two centers. Longitudinal data from the measurement at baseline (T0) and after 1, 2 and 5 years (T1, T2 and T5, respectively) follow-up were analyzed.

Participants

Inclusion criteria for the CHECK study were age 45–64 year, hip and/or knee pain for which the participant not yet had consulted their physician, or the first consultation was within 6 months before entry. Participants were excluded if hip or knee pain was based on any other pathological condition that could explain the symptoms (for example the presence of inflammatory rheumatic disorders, joint prosthesis hip and knee, previous joint trauma and serious co morbidity). All participants provided written informed consent before entering the study, and the Medical Ethical Board of hospital ‘Medisch Spectrum Twente’ in Enschede, the Netherlands, approved the study. Participants from two of the 10 CHECK centers were invited to participate in this study on work capacity.

Procedure

From the before mentioned regional CHECK study centers of Groningen and Enschede all 157 participants were invited to participate in this longitudinal (5 years) study on performance based measurement of physical capacity. A few days after each measurement participants were telephoned about adverse effects, muscle soreness, and increases of symptoms due to the physical tests.

Measurements

Functional Capacity the WorkWell Systems Functional Capacity Evaluation (formerly known as Isernhagen Work Systems FCE; hereafter: FCE) [WorkWell Systems 2006] was used to assess participants’ capacity to perform work-related activities. Twenty-two tests, including all those that cause load bearing to the hips and the knees, were selected from the 2-day FCE protocol at baseline (T0); at T1, T2 and T5 the tests were performed on 1 day. These tests aim to record capacity with regards to Manual Material Handling, Postural Tolerance and Repetitive Movements, and refer to physical strength, endurance or speed. Providing the evaluator judged the tests to be performed safely (based on observation criteria as movement pattern and postural changes) participants were asked to continue to a higher load level (5 repetitions per level). The static endurance tests were continued until a pre-set limit was reached. The participant was

free to end any test at any moment, for example because of discomfort or pain. Participants were informed about the possibility of muscle soreness and increase of symptoms after maximal testing and the natural course of this [25].

The longitudinal analyses were performed on the results from 6 of the 22 tests that represent main aspects of physical job demands and the baseline data of which were presented earlier in a cross-sectional study on comparison with healthy workers [19]. The reliability of these tests has been established [26, 27]. Tests were performed on 2 consecutive days at baseline (with repeat of material handling tests on the second day), and on 1 day in all other occasions.

Manual Material Handling was tested in the following FCE tests:

Lifting Low

Objective: capacity of lifting from table to floor. Materials: plastic receptacle (40×30×26 cm), a wall mounted system with adjustable shelves and weights of 1.0, 2.0 and 4.0 kg. Procedure: five lifts from table at 74 cm to floor and vice versa in standing position within 90 s. Four to five weight increments until maximum amount of kg was reached.

Lifting High

Objective: capacity of overhead lifting task. Materials: plastic receptacle (40×30×26 cm), a wall mounted system with adjustable shelves and weights of 1.0, 2.0 and 4.0 kg. Procedure: five lifts from table (74 cm) to crown height and vice versa in standing position within 90 s. Four to five weight increments until maximum amount of kg was reached.

Carrying Long

Objective: capacity of two handed carrying. Materials: plastic receptacle (40×30×26 cm), a wall mounted system with adjustable shelves and weights of 1.0, 2.0 and 4.0 kg. Procedure: 20 m carrying at waist height with receptacle within 90 s. Four to five weight increments until maximum amount of kg was reached.

Postural tolerance was tested in the following FCE tests:

Overhead Working

Objective: capacity of postural tolerance of overhead working. Materials: aluminium plate adjustable in height with 20 holes, bolts and nuts and two cuff weights of 1.0 kg each. Procedure: standing with hands at crown height, manipulating nuts and bolts wearing cuff weights around the wrists. The time that position is held was measured (sec).

Repetitive movements were tested in the following FCE tests:

Repetitive Bending

Objective: capacity of repetitive bending and reaching. Materials: 20 marbles and 2 bowls with a 14-cm diameter positioned at floor and crown height. Procedure: standing with knees flexed between 0° and 30°, move marbles vertically from floor to crown height as fast as possible. Time needed to remove 20 marbles is scored (sec).

Repetitive Rotations

Objective: capacity of fast repetitive side movements of the upper extremity. Materials: 30 marbles and 2 bowls with a 14-cm diameter positioned at table height (74 cm). Procedure: sitting with bowls on wingspan distance, move marbles horizontally at table height from right to left with right arm as fast as possible and vice versa. Time needed to move 30 marbles is scored (sec).

Secondary Measures

Preceding the FCE tests participants' age and sex were registered. Body height and weight were measured and Body Mass Index (BMI) was calculated. Tests were administered by 4th year physical therapy students who had received a 1-day training in the procedures and the execution of the FCE. They were trained and supervised by the research team.

Data on work status (either or not having a paid job for 12 h per week) and self-reported function (Western Ontario McMasters Arthritis Scale, WOMAC, scale 0–100 [28]) were retrieved from the questionnaires the participants filled out in the CHECK study. Radiographic data on the stage of OA (classified as Kellgren and Lawrence (K&L) rating score [29]) were retrieved from measurements at baseline, T2 and T5 to document the radiographic damage of the joints. Clinical diagnosis of OA was classified according to the clinical American College of Rheumatology (ACR) criteria for OA [30, 31].

Statistical Analysis

For identifying independent prognostic factors for performance on each of the 6 selected FCE tests the variables sex, age, having a paid job, BMI, WOMAC function and time of measurement (Tx) were entered in the multilevel analysis (IBM SPSS Statistics 20). These independent variables were entered, using the following cut-off points: age categories 45–49, 50–54, 55–59, ≥ 60; BMI: <18.5 (underweight), 18.5–24.9 (normal weight), 25.0–29.9 (overweight), ≥ 30 (obese); WOMAC < 69 points and ≥ 69 points, based on a

previous study [32]; $T_x=0, 1, 2, 5$. We used multilevel analysis estimated fixed effects, because measurement moments were nested within participants.

To identify subgroups of participants with possibly different courses of functional capacity in time, Latent Class Analyses (LatentGold 4.5) were performed. The number of classes that best fitted the data was determined using the Bayesian Information Criterion (BIC). The identified classes were related to sex, age, having a paid job, BMI, and WOMAC function.

To analyze for selective drop-out, the participants who participated only once and those who participated more than once were compared for differences in performance at baseline. To check for the influence of missing data, results were compared to those in an imputed set of data. For missing values, multiple imputation was used (from the statistical package SPSS) with the ‘predictive mean matching’ model type, choosing to create 20 imputed datasets in this study. We included covariates, outcome (FCE) and auxiliary variables that help to impute variables with missing data.

Results

Ninety-six of the 157 participants (61%) from two of the CHECK centers participated in the functional capacity evaluations. The baseline characteristics of the 96 participants (Table 1) showed no significant differences with the 907 others in the cohort. The majority were female, about 50% had a paid job for 12 h or more per week.

At inclusion CHECK was indeed an early OA cohort, the proportion of participants with radiological osteoarthritis ($K\&L > 1$) was 6% for the knee and 10% for the hip; however, 76% of the patients with knee symptoms could be diagnosed as OA according to the clinical American College of Rheumatology (ACR) criteria. A quarter of the 1003 CHECK participants with hip symptoms (24%) fulfilled the clinical classification criteria of hip OA. Table 2

demonstrates that our study sample ($n=96$) showed no $K\&L$ scores > 1 at baseline (data from 5 to 6 participants were missing) and relatively more knee than hip OA at follow-up.

During this longitudinal study, the participating sample decreased from $n=96$ at baseline to $n=64$ (67%) after 1 year, $n=61$ (64%) after 2 years and $n=35$ (36%) after 5 years. Reasons to drop out that were given in the telephone interviews were the physical load of the tests, physical symptoms in the days following the test, serious health problems unrelated to OA, personal circumstances, moved to other part of the country, or stopped participation in CHECK.

Course of the Performance

Overall there was no change in the results over the 5 years (Fig. 1), on average the lines were more or less flat or showed no clear trend (Fig. 1). For example, in the test Lifting Low participants lifted on average around 25 kg at baseline (T_0) and this performance remained stable over the 5 years. Because of this flat course of the functional capacity, the factors in the regression model (Table 3) can be considered related to the average performance.

Female sex and a low WOMAC score both were statistically significant co-variables for poorer performance on the three manual material handling tests (lifting low, lifting high, carrying; Table 3).

On lifting low for example, females lifted on average 15.8 kg. less than males and participants with a score lower than 69 point on WOMAC function lifted 2.78 kg. less than those with score of 69 or higher. Older age was a statistically significant prognostic factor for lower performance on three tests (two manual material handling tests, for example weight lifted on lifting high was 0.19 kg. lower per year older, and one repetitive movements test). It took obese participants about 5 s longer than the participants with normal weight and overweight to complete the repetitive knee bending test. Having a paid job was not related to any of the six tests. On all comparisons of results on the 6 FCE test

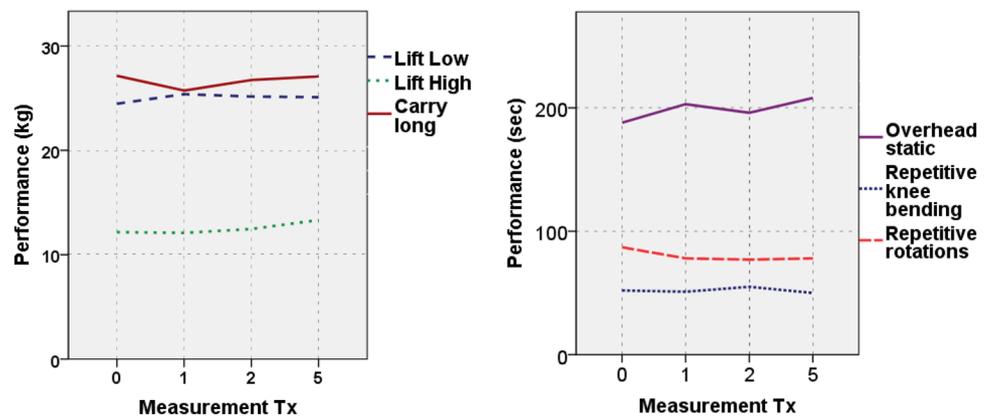
Table 1 Subject characteristics at baseline in the cohort ($n=1003$)

	Participants FCE ($n=96$)	Others in cohort ($n=907$)	Mean	CI 95% of the difference		P
	Mean (SD)/or %		Difference	Lowerbound	Upperbound	
Age (years)	56 (4.9)	56 (5.2)	0	-1.09	1.09	0.995
Female (%)	84%	78%	6%	-0.13	0.02	0.173
Work participation (%)	48%	46%	2%	-0.12	0.09	0.651
BMI (kg/m^2)	26.0 (4.5)	26.2 (4.1)	-0.2	-1.07	0.67	0.845
WOMAC						
Function	73.4 (17.8)	76.9 (17.1)	-3.6	-7.10	0.10	0.094
Stiffness	65.3 (21.9)	67.0 (21.0)	-1.7	-6.14	2.74	0.435
Pain	73.0 (17.1)	74.9 (17.1)	-1.9	-5.50	1.70	0.290

Table 2 Radiographic progression of OA in the FCE participants: baseline, 2 and 5 years: frequencies of Kellgren and Lawrence scores

Joint	K&L grade	Measurement Tx =			Joint	K&L grade	Measurement Tx =		
		0	2	5			0	2	5
Hip, left	0	68	59	55	Hip, right	0	69	62	61
	1	22	30	32		1	22	26	25
	2	0	1	3		2	0	2	4
	3	0	0	0		3	0	0	0
	4	0	0	0		4	0	0	0
	N =	90	89	90		N =	91	90	90
Knee, left	0	41	36	25	Knee, right	0	43	37	30
	1	50	44	45		1	47	42	43
	2	0	12	19		2	0	12	15
	3	0	1	3		3	0	1	4
	4	0	0	0		4	0	0	0
	N =	91	93	92		N =	90	92	92

Fig. 1 Estimated marginal means for the strength tests (left) and for the endurance and speed tests (right)



between different measurements in time (Tx) there were only two significant differences (between baseline and 5-years follow-up measurement for overhead static and repetitive rotation tests). Results for the three manual material handling tests (lifting low and high, carrying) were very consistent, those for the endurance (static overhead) and speed tests (repetitive bending and repetitive rotation) were not.

The objective of identifying classes of participants with different courses (trajectories) of functional capacity over time, was answered by using the latent class analysis. Based on the statistical Bayesian Information Criterion (BIC; Supplementary Table 4) models with three latent classes were selected for the three manual material handling tests and for the overhead static posture test (Figs. 2, 3); on the latter test BIC indicated 4 classes, but because of the small difference with a model with 3 classes and considering the sample size a three classes model was selected. For the repetitive movements test, models with two classes were selected (not shown for reasons of space). The main difference between the trajectories was the level of performance, while the course over time (T0–T1–T2–T5) was not very different. For

the lifting and carrying tests we identified three subgroups and labeled them as ‘weak giving way’, ‘stable and able’ and ‘strong with decline’. The composition of the identified classes on the manual material handling test was compared with regards to the personal characteristics (sex, BMI, WOMAC and work status). This comparison showed a consistent increase (from lower to higher lines in Figs. 2, 3) in the proportions of males, participants with higher WOMAC scores and overweight participants (tables in Figs. 2, 3). For example, the left columns of the table in Fig. 2 show that on the test Lifting Low the proportion of participants with WOMAC score ≥ 69 increased from 44% in the low class to 75% in the medium and 83% in the high class. On the static endurance and repetitive movements tests the picture is less consistent (latter not shown).

Influence of Missing Data

Multiple imputation resulted in a model with two instead of three latent classes for the FCE-test carrying. For the other five tests similar models as without imputation could be

Table 3 Predictors of performance on the 6 FCE tests (multilevel analysis, statistically significant factors printed bold)

Parameter	Model											
	Estimate	p-value										
Intercept	34.4	<0.001	18.4	<0.001	37.39	<0.001	249.82	<0.001	50.18	<0.001	76.84	<0.001
Female (vs. male)	-15.8	<0.001	-8.79	<0.001	-18.43	<0.001	-49.54	0.02	4.62	0.065	-2.77	0.421
Norm weight vs. obese	0.75	0.60	0.8	0.29	2.27	0.146	-1.44	0.93	-5.35	0.015	-1.02	0.737
Over weight vs. obese	1.56	0.26	1.71	0.02	3.52	0.023	11.71	0.41	-5.53	0.015	-2.72	0.382
Paid work (vs. no work)	-0.2	0.85	0.04	0.95	-0.51	0.677	-18.21	0.08	-3.16	0.079	0.67	0.784
Age (centred at 55)	-0.22	0.08	-0.19	0.005	-0.43	0.002	0.43	0.78	0.52	0.006	0.39	0.146
Tx=0 (vs. Tx=5)	-0.62	0.63	-1.16	0.09	0.07	0.965	-27.28	0.03	1.83	0.444	10.31	0.002
Tx=1 (vs. Tx=5)	0.29	0.82	-1.23	0.08	-1.35	0.38	-12.08	0.31	0.27	0.914	0.11	0.973
Tx=2 (vs. Tx=5)	0.06	0.96	-0.86	0.21	-0.34	0.823	-19.02	0.1	4.07	0.105	0.15	0.964
WOMAC < 69 (vs. ≥ 69)	-2.78	0.003	-1.96	<0.001	-3.01	0.006	9.5	0.3	3.06	0.062	3.5	0.119

generated, and for the 5-year measurements the confidence intervals were often relatively wide (not shown).

Those who participated only once did not show significantly different performances on the FCE compared to those who participated more than once (results not shown), indicating that there was no selective drop-out of participants with lower functional capacity at baseline.

Discussion

Main Results

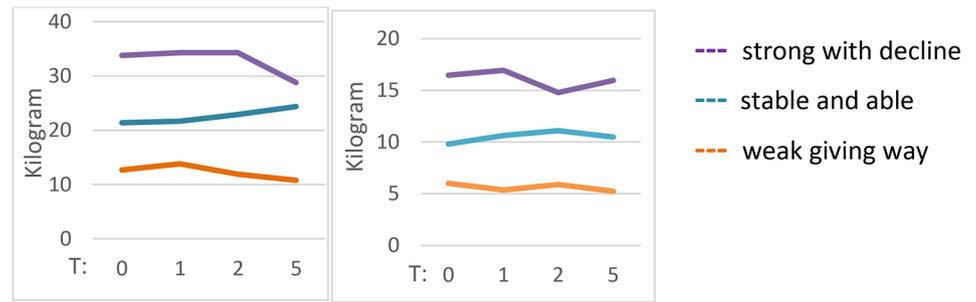
The functional capacity for manual material handling, static overhead work and repetitive movements showed hardly any change over 5 years. The strongest independent co-variables of a higher functional capacity were male sex and a high self-reported physical function (WOMAC ≥ 69); lower age was a significant factor on three tests. Work status (either or not having a paid job) and progression of time were unrelated to functional capacity. For manual material handling capacity three trajectories (classes of participants) could be identified, a very strong class with mostly males (‘strong with decline’), a group with medium strong participants (males and females, ‘stable and able’) and a weak group (only females, ‘weak giving way’). All three groups showed little or no change over time.

Methodology

A strength of our study is that it is the first longitudinal study on performance based measurement of functional capacity. In line with the objectives of the CHECK study, it gives basic insights into the course of functional capacity in OA over time. Measuring performance may offer additional information on outcomes to self-reported questionnaire data when it concerns work ability [32]. To allow optimal use of all available data while dealing with drop out, multi-level analysis was used. With this method all data could be used and participants who completed more measurements contributed more to the estimations.

A weakness of the study is that many participants did not complete all follow-up measurements, especially in the 5-years measurement there was loss to follow-up which may have introduced selection bias. Multiple imputations did not lead to rejection of the identified classes, except for the FCE-test ‘carrying’. The method we chose for dealing with missing data in the analytic phase was multiple imputation. Multiple imputation offers some protection against a simple missing not at random (MNAR) mechanism [33]. This method is more likely to be valid and should be preferred because of its greater efficiency, because multiple imputation may be more robust to departures from assumptions of

Fig. 2 Graphs depict trajectories of performance on the FCE lifting tests for the three latent classes over measurements (T=0, 1, 2, 5). Percentages below graphs indicate proportions of subjects in each class, from lowest to upper curve (sum=100%). Tables present composition of classes by personal characteristics (proportions for sex, BMI, WOMAC score, work status) per class (vertical sum=100% for each variable)

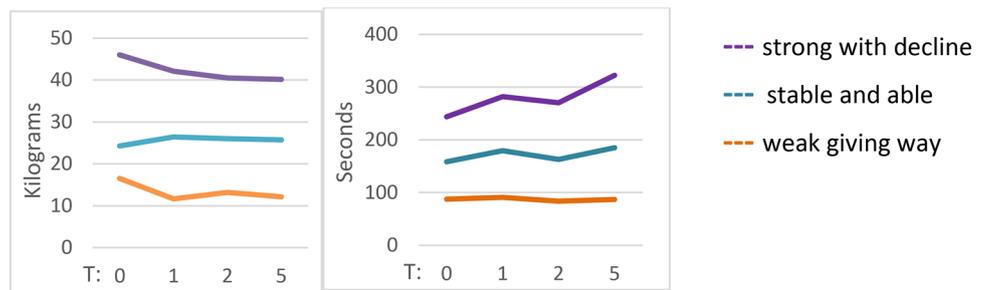


(a) Lifting low (41–44–15%). (b) Lifting high (41–36–23%).

Lift Low, % within class				Lift high, % within class,		
lower	medium	upper		lower	medium	upper
100	92	25	Female	100	95	42
0	8	75	vs male	0	5	58
52	41	36	BMI norm	49	43	34
25	34	51	vs over	26	33	49
23	24	14	vs obese	25	24	17
56	25	17	WM<69	55	30	14
44	75	83	vs >=69	45	70	86
25	50	54	Work yes	23	52	48
75	50	46	vs no	77	48	52

WM=WOMAC

Fig. 3 Graphs depict trajectories of performance on the FCE tests carrying (kg) and overhead work (seconds) for the 3 latent classes over measurements (T=0, 1, 2, 5). Percentages below graphs indicate proportions of subjects in each class, from lowest to upper curve (sum=100%). Tables present composition of classes by personal characteristics (proportions for sex, BMI, WOMAC score, work status) per class (vertical sum=100% for each variable)



(a) Carrying (49–43–8%). (b) Overhead static work (18–54–28%).

Carrying, % within class,				Static work, % within class,		
lower	medium	upper		lower	medium	upper
100	84	3	Female	94	90	71
0	16	97	vs male	6	10	29
52	36	43	BMI norm	39	49	38
24	40	51	vs over	32	28	47
24	24	7	vs obese	29	23	15
52	22	13	WM<69	64	35	16
48	78	87	vs >=69	36	65	84
27	48	66	Work yes	43	37	45
73	52	34	vs no	57	63	55

WM=WOMAC

worst-case and best-case scenarios. The advantage of this method is that it represents the uncertainty about which value to impute. Sensitivity analyses incorporating worst-case and best-case scenarios is one of the several statistical approaches for dealing with missing data. Because the limitations of too small standard error and thereby overestimation of precision of the results that were reported in the literature [34], we chose for the multiple imputation as method for dealing with the missing data. Proving that the data is missing at random (MAR) is very hard, but it is likely that the MAR assumption is more plausible when more variables are included in the multiple imputation model. The argument that supports the choice for MAR was based on the comparison between persons who were present at all three measuring moments (complete cases) and persons who were absent for at least one measurement moment at baseline (see Online Appendix Table A02). In addition, to give more insight into the imputation, the imputed data is presented both from baseline and the last measurement moment (Online Appendix Tables A02 and A03).

Participants who dropped out after the first measurement did not perform different than the others at baseline. In phone contacts following the tests several reasons for drop-out were recorded, varying from having moved to another part of the country to not wanting to take time off from work to participate in the tests. Also some of the participants indicated they did not want to continue because of the physical strenuousness of the tests and subsequent muscle soreness or increased musculoskeletal pain. Not their capacity, but differences in tolerance to discomfort or pain after the tests, may have influenced participants' decision to stop participating in the study. Pain acceptance, pain self-efficacy and perceived work load have been shown relevant factors for staying at work [35]. Although the FCE's in this study were not performed in an occupational health context, the relation between pain and performance during the FCE is a subject for further study that may yield interesting information for occupational and general health professionals.

Some of the results support the validity of FCE tests: as expected [18], men demonstrated higher manual material handling performances than women. Another valid results was that tests stressing the hips and/or the knees (for example lifting low and carrying) showed smaller proportions of participants in the strongest classes (15% and 8% respectively) compared to the other tests. The three manual material handling tests showed very consistent results, whereas the static endurance and repetitive speed tests gave more varying results. Work status (having a paid job) was not consistently related to FCE performance, perhaps because physical demands in a job were not evaluated and may have been different between participants. For example, the performance on lifting high may be dependent on exposure to (and implicating training of) this task at work; working

participants with office jobs are not likely to lift high loads on this test. The composition of the resulting classes has been influenced by the composition of the study population (86% females).

Clinical Relevance and Implications

Our finding of stable functional capacities over a 5 years period is different from 'general opinions' about the effects of both aging and OA on physical functioning. This observation demonstrates that work-related functional capacity does not necessarily decline in 5 years in participants aged in their fifties, with early OA. Interestingly, Holla et al. [9] also identified three subgroups in the CHECK study; however, they included only the participants with knee OA and used self-reported function (WOMAC) as the outcome measure to identify subgroups of participants with good, moderate or poor outcome of their knee disorder. Similar to our study, which evaluated physical function with performance based testing, Holla's subgroups demonstrated different levels of function and the changes in function over time were small [9]. Considering the variation between participants with early OA of hips and knees generalization is difficult, but identifying and describing subgroups may be an important step in the clinical management of functioning problems in OA.

Except perhaps for jobs with a high physical load, the consequences of aging (and OA) on work capacity should not be overestimated. Discrepancies between work capacity and work load need to be studied on the level of the individual, including changes over time and the effect of interventions aiming at these discrepancies. In a previous study in the whole cohort (N = 1002) 20% of the participants reported they had made work accommodations and 26% would like to do so. Decreased working hours were most frequently reported, besides adaptations in the work place and the work technique. Workplace interventions have been shown to be effective for individuals with OA [36] and may help to cope with OA related work difficulties. Physiotherapy may be an alternative or additional intervention to maintain physical capacity. The indication for either of these interventions may be supported by knowledge from an FCE.

Maybe the participation in a cohort study worked as a stimulus, as some participants indicated that the information they received during the study made them more aware of factors that could influence their health situation and functioning. Changes in life-style, for example starting or continuing regular physical activity and training, or trying to keep at a constant body weight, could contribute to maintaining functional capacity with climbing age [21, 37]. Health information in an early stage of OA may have a preventive effect on physical function.

Although increasing age (as within factor) had no effect in this study, we did find effects of age differences (between factor) on physical performance. This seems contradictory, but may be due to bias by many possible factors; it could be a cohort effect in which the younger participants have benefited from better health care in any period in their life or from improved work conditions; also the older participants probably have a longer work history and have therefore been longer exposed to physical load [38].

The quantification of functional capacity by performance based testing may be helpful as a reference for assessments and interventions in the field of occupational health, specifically with an ageing population that needs to postpone retirement. The identification of the class ‘stable and able’ regarding performance on the lifting low test is interesting. Their performance level is around 23 kg, which is equal to the maximally recommended weight to lift in ideal working situations (NIOSH, [39]) and to the lower limit of the category of functions with a medium physical work load according to the DOT [40]. This means that there is a category of workers (males and females) for whom an FCE may be an indicated instrument to monitor the match between work load and physical capacity of the individual. It has to be noticed that the limits of agreement of these measurements are wide [27], which makes clinical interpretation of individual results challenging.

Both work load (job characteristics) and personal capacity (maintaining good physical function) may affect workers in maintaining a sustainable working life. The conclusion of our study is that in subgroups of participants with early symptomatic OA, determined by age, sex and self-reported function, physical work capacity seems to be a stable characteristic over 5 years.

Acknowledgements CHECK-cohort funded by the Dutch Arthritis Foundation on the lead of a steering committee comprising 16 members with expertise in different fields of OA chaired by Prof. Dr. J. W. J. Bijlsma and coordinated by J. Wesseling. Involved are: Erasmus Medical Center Rotterdam; Kennemer Gasthuis Haarlem; Leiden University Medical Center; Maastricht University Medical Center; Martini Hospital Groningen/Allied Health Care Center for Rheum and Rehabilitation Groningen; Medical Spectrum Twente Enschede/Ziekenhuisgroep Twente Almelo; Reade, formerly Jan van Breemen Institute/VU Medical Center Amsterdam; St. Maartens-kliniek Nijmegen; University Medical Center Utrecht and Wilhelmina Hospital Assen.

Author Contributions All authors have made substantial contributions to the conception and design of the study, or acquisition of data, or analysis and interpretation of data, drafting the article or revising it critically for important intellectual content and final approval of the version to be submitted. BHJ: Conception and design, Analysis and interpretation of the data, Drafting of the article, Critical revision of the article for important intellectual content, Final approval of the article, Provision of study materials or patients, Statistical expertise, Obtaining of funding and Collection and assembly of data. SR: Analysis and interpretation of the data, Critical revision of the article for important intellectual content, Final approval of the article, Statistical expertise.

RMF: Conception and design, Analysis and interpretation of the data, Drafting of the article, Critical revision of the article for important intellectual content, Final approval of the article. IWM: Conception and design, Analysis and interpretation of the data, Critical revision of the article for important intellectual content, Final approval of the article, Provision of study materials or patients. SCP: Critical revision of the article for important intellectual content, Final approval of the article, Collection and assembly of data. D-BKW: Conception and design, Critical revision of the article for important intellectual content, Final approval of the article, Provision of study materials or patients. OFGJ: Conception and design, Drafting of the article, Critical revision of the article for important intellectual content, Final approval of the article, Obtaining of funding. HJB and FGJO take responsibility for the integrity of the work as a whole.

Funding The CHECK cohort study was funded by the Dutch Arthritis Foundation (Reumafonds).

Compliance with Ethical Standards

Conflict of interest None of the authors declare any conflicts of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

- Vignon E, Valat J-P, Rosignol M, Avouac B, Rozenberg S, Thoumie P, et al. Osteoarthritis of the knee and hip and activity: a systematic international review and synthesis (OASIS). *Joint Bone Spine Rev Rhum*. 2006;73(4):442–455.
- Palmer KT. The older worker with osteoarthritis of the knee. *Br Med Bull*. 2012;102(1):79–88.
- Agaliotis M, Mackey MG, Jan S, Fransen M. Burden of reduced work productivity among people with chronic knee pain: a systematic review. *Occup Environ Med*. 2014;71(9):651–659.
- Bieleman HJ, Bierma-Zeinstra SMA, Oosterveld FGJ, Reneman MF, Verhagen AP, Groothoff JW. The effect of osteoarthritis of the hip or knee on work participation. *J Rheumatol*. 2011;38(9):1835–1843.
- Sharif B, Garner R, Hennessy D, Sanmartin C, Flanagan WM, Marshall DA. Productivity costs of work loss associated with osteoarthritis in Canada from 2010 to 2031. *Osteoarthritis Cartilage*. 2017;25(2):249–258.
- <https://www.volksgezondheidszorg.info/onderwerp/artrose/cijfers-context/huidige-situatie#node-prevalentie-artrose-naar-leeftijd-en-geslacht>. Gommer AM, Poos MJJC, Hulshof T, editors Visited June 22nd 2018.
- Douwes M, van Genabeek J, van den Bossche S, editors. *Arbobalans 2016. Kwaliteit van de arbeid, effecten en maatregelen in Nederland (Quality of work, effects and measures in the Netherlands)*. Leiden: TNO; 2016.
- van Dijk GM, Veenhof C, Spreeuwenberg P, Coene N, Burger BJ, van Schaardenburg D, et al. Prognosis of limitations in activities in osteoarthritis of the hip or knee: a 3-year cohort study. *Arch Phys Med Rehabil*. 2010;91(1):58–66.

9. Holla JFM, van der Leeden M, Heymans MW, Roorda LD, Bierma-Zeinstra SMA, Boers M, et al. Three trajectories of activity limitations in early symptomatic knee osteoarthritis: a 5-year follow-up study. *Ann Rheum Dis*. 2014;73(7):1369–1375.
10. Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. *Osteoarthritis Cartilage*. 2014;22(12):2041–2050.
11. Dowsey MM, Smith AJ, Choong PFM. Latent class growth analysis predicts long term pain and function trajectories in total knee arthroplasty: a study of 689 patients. *Osteoarthritis Cartilage*. 2015;23(12):2141–2149.
12. Wilkie R, Blagojevic-Bucknall M, Jordan KP, Pransky G. Onset of work restriction in employed adults with lower limb joint pain: individual factors and area-level socioeconomic conditions. *J Occup Rehabil*. 2013;23(2):180–188.
13. Wilkie R, Phillipson C, Hay E, Pransky G. Frequency and predictors of premature work loss in primary care consultants for osteoarthritis: prospective cohort study. *Rheumatol Oxf Engl*. 2014;53(3):459–464.
14. Takala E-P, Pehkonen I, Forsman M, Hansson G-A, Mathiassen SE, Neumann WP, et al. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scand J Work Environ Health*. 2010;36(1):3–24.
15. James CL, Reneman MF, Gross DP. Functional capacity evaluation research: report from the second international functional capacity evaluation research meeting. *J Occup Rehabil*. 2016;26(1):80–83.
16. Hamberg-van Reenen HH, Ariëns GAM, Blatter BM, van Mechelen W, Bongers PM. A systematic review of the relation between physical capacity and future low back and neck/shoulder pain. *Pain*. 2007;130(1–2):93–107.
17. Bugajska J, Makowiec-Dąbrowska T, Bortkiewicz A, Gadzicka E, Marszałek A, Lewandowski Z, et al. Physical capacity of occupationally active population and capability to perform physical work. *Int J Occup Saf Ergon*. 2011;17(2):129–138.
18. Soer R, van der Schans CP, Geertzen JH, Groothoff JW, Brouwer S, Dijkstra PU, et al. Normative values for a functional capacity evaluation. *Arch Phys Med Rehabil*. 2009;90(10):1785–1794.
19. Bieleman HJ, van Ittersum MW, Groothoff JW, Oostveen JCM, Oosterveld FGJ, van der Schans CP, et al. Functional capacity of people with early osteoarthritis: a comparison between subjects from the cohort hip and cohort knee (CHECK) and healthy ageing workers. *Int Arch Occup Environ Health*. 2010;83(8):913–921.
20. Savinainen M, Nygård C-H, Ilmarinen J. A 16-year follow-up study of physical capacity in relation to perceived workload among ageing employees. *Ergonomics*. 2004;47(10):1087–1102.
21. Kenny GP, Yardley JE, Martineau L, Jay O. Physical work capacity in older adults: implications for the aging worker. *Am J Ind Med*. 2008;51(8):610–625.
22. de Zwart BC, Broersen JP, Frings-Dresen MH, van Dijk FJ. Repeated survey on changes in musculoskeletal complaints relative to age and work demands. *Occup Environ Med*. 1997;54(11):793–799.
23. Soer R, Brouwer S, Geertzen JH, van der Schans CP, Groothoff JW, Reneman MF. Decline of functional capacity in healthy aging workers. *Arch Phys Med Rehabil*. 2012;93(12):2326–2332.
24. Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SMA, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the osteoarthritis initiative. *Ann Rheum Dis*. 2009;68(9):1413–1419.
25. Soer R, Groothoff JW, Geertzen JHB, van der Schans CP, Reesink DD, Reneman MF. Pain response of healthy workers following a functional capacity evaluation and implications for clinical interpretation. *J Occup Rehabil*. 2008;18(3):290–298.
26. Gross DP, Battié MC. Reliability of safe maximum lifting determinations of a functional capacity evaluation. *Phys Ther*. 2002;82(4):364–371.
27. van Ittersum MW, Bieleman HJ, Reneman MF, Oosterveld FGJ, Groothoff JW, van der Schans CP. Functional capacity evaluation in subjects with early osteoarthritis of hip and/or knee; is two-day testing needed? *J Occup Rehabil*. 2009;19(3):238–244.
28. Roorda LD, Jones CA, Waltz M, Lankhorst GJ, Bouter LM, van der Eijken JW, et al. Satisfactory cross cultural equivalence of the Dutch WOMAC in patients with hip osteoarthritis waiting for arthroplasty. *Ann Rheum Dis*. 2004;63(1):36–42.
29. Kellgren JH, Lawrence JS. Radiological assessment of rheumatoid arthritis. *Ann Rheum Dis*. 1957;16(4):485–493.
30. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum*. 1991;34(5):505–514.
31. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and therapeutic criteria committee of the American Rheumatism association. *Arthritis Rheum*. 1986;29(8):1039–1049.
32. Bieleman HJ, Reneman MF, van Ittersum MW, van der Schans CP, Groothoff JW, Oosterveld FGJ. Self-reported functional status as predictor of observed functional capacity in subjects with early osteoarthritis of the hip and knee: a diagnostic study in the CHECK cohort. *J Occup Rehabil*. 2009;19(4):345–353.
33. Kontopantelis E, White IR, Sperrin M, Buchan I. Outcome-sensitive multiple imputation: a simulation study. *BMC Med Res Methodol*. 2017;17(1):2.
34. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, Kristensen NR, Pham TM, Pedersen L, et al. Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol*. 2017;9:157–166.
35. de Vries HJ, Reneman MF, Groothoff JW, Geertzen JHB, Brouwer S. Self-reported work ability and work performance in workers with chronic nonspecific musculoskeletal pain. *J Occup Rehabil*. 2013;23(1):1–10.
36. Baldwin D, Johnstone B, Ge B, Hewett J, Smith M, Sharp G. Randomized prospective study of a work place ergonomic intervention for individuals with rheumatoid arthritis and osteoarthritis. *Arthritis Care Res*. 2012;64(10):1527–1535.
37. Jackson AS, Sui X, Hébert JR, Church TS, Blair SN. Role of lifestyle and aging on the longitudinal change in cardiorespiratory fitness. *Arch Intern Med*. 2009;169(19):1781–1787.
38. Christensen K, Thinggaard M, Oksuzyan A, Steenstrup T, Andersen-Ranberg K, Jeune B, et al. Physical and cognitive functioning of people older than 90 years: a comparison of two Danish cohorts born 10 years apart. *Lancet*. 2013;382(9903):1507–1513.
39. Waters TR, Baron SL, Piacitelli LA, Anderson VP, Skov T, Haring-Sweeney M, et al. Evaluation of the revised NIOSH lifting equation. A cross-sectional epidemiologic study. *Spine*. 1999;24(4):386–394.
40. Reneman MF, Kuijjer W, Brouwer S. Item validity of the physical demands from the dictionary of occupational titles for FCE of clients with chronic back pain. *Work*. 2006;26(2):229 (author reply 231–233).