



University of Groningen

Association between social factors and performance during Functional Capacity Evaluations

Ansuategui Echeita, Jone; van Holland, Berry J; Gross, Douglas P; Kool, Jan; Oesch, Peter; Trippolini, Maurizio A; Reneman, Michiel F

Published in:
Disability and Rehabilitation

DOI:
[10.1080/09638288.2018.1448120](https://doi.org/10.1080/09638288.2018.1448120)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Ansuategui Echeita, J., van Holland, B. J., Gross, D. P., Kool, J., Oesch, P., Trippolini, M. A., & Reneman, M. F. (2019). Association between social factors and performance during Functional Capacity Evaluations: a systematic review. *Disability and Rehabilitation*, 41(16), 1863-1873.
<https://doi.org/10.1080/09638288.2018.1448120>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Association between social factors and performance during Functional Capacity Evaluations: a systematic review

Jone Ansuategui Echeita, Berry J. van Holland, Douglas P. Gross, Jan Kool, Peter Oesch, Maurizio A. Trippolini & Michiel F. Reneman

To cite this article: Jone Ansuategui Echeita, Berry J. van Holland, Douglas P. Gross, Jan Kool, Peter Oesch, Maurizio A. Trippolini & Michiel F. Reneman (2018): Association between social factors and performance during Functional Capacity Evaluations: a systematic review, *Disability and Rehabilitation*, DOI: [10.1080/09638288.2018.1448120](https://doi.org/10.1080/09638288.2018.1448120)

To link to this article: <https://doi.org/10.1080/09638288.2018.1448120>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 09 Mar 2018.



[Submit your article to this journal](#)



Article views: 49



[View related articles](#)



[View Crossmark data](#)

Association between social factors and performance during Functional Capacity Evaluations: a systematic review

Jone Ansuategui Echeita^a, Berry J. van Holland^b, Douglas P. Gross^c, Jan Kool^d, Peter Oesch^d, Maurizio A. Trippolini^{e,f,g} and Michiel F. Reneman^a

^aDepartment of Rehabilitation Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands;

^bInstitute for Sports Studies, Hanze University of Applied Sciences, Groningen, The Netherlands; ^cDepartment of Physical Therapy, University of Alberta, Edmonton, Canada; ^dRehabilitation Centre Valens, Valens, Switzerland; ^eCenter for Disability Research, Liberty Mutual Research Institute for Safety, Boston, MA, USA; ^fMassachusetts General Hospital, Institute for Health Professions, Boston, MA, USA; ^gDepartment of Work Rehabilitation, Rehaklinik Bellikon, Bellikon, Switzerland

ABSTRACT

Purpose: Determine the association of different social factors with Functional Capacity Evaluation (FCE) performance in adults.

Materials and methods: A systematic literature search was performed in MEDLINE, CINAHL, and PsycINFO electronic databases. Studies were eligible if they studied social factor's association with the performance of adults undergoing FCE. Studies were assessed on methodological quality and quality of evidence. The review was performed using best-evidence synthesis methods.

Results: Thirteen studies were eligible and 11 social factors were studied. Considerable heterogeneity regarding measurements, populations, and methods existed among the studies. High quality of evidence was found for the association of FCE performance with the country of FCE and examiner's fear behavior; moderate quality of evidence with previous job salary; and low or very low quality of evidence with compensation status, litigation status, type of instruction, time of day (workday), primary or mother language, and ethnicity. Other social factors were not studied.

Conclusions: Evidence for associations of various social factors with FCE performance was found, but robust conclusions about the strength of the associations cannot be made. Quality of evidence ranged from high to very low. Further research on social factors, also within a biopsychosocial context, is necessary to provide a better understanding of FCE performance.

ARTICLE HISTORY

Received 14 July 2017
Revised 28 February 2018
Accepted 1 March 2018

KEYWORDS

Functional Capacity Evaluation; occupational rehabilitation; performance assessment; environmental factors; systematic review

► IMPLICATIONS FOR REHABILITATION


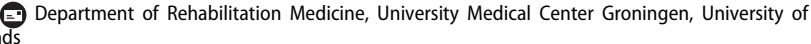
- Research on Functional Capacity Evaluation (FCE) performance and its association with biopsychosocial factors have scarcely addressed the impact of social factors, limiting full understanding of FCE results.
- The social factors, healthcare (examiner's fear behavior and type of instruction), personal or cultural systems (country of FCE, primary or mother language, and ethnicity), workplace system (previous job salary, time of day (workday)), and legislative and insurance system (compensation and litigation status), have a bearing in FCE performance.
- Better understanding of factors associating with functional capacity provide insights in FCE, allowing clinicians to improve the evaluations and interpretations of the assessment and better design the rehabilitation program.
- Better understanding of factors that influence FCE performance, and of unstudied factors, will allow researchers guidance to further investigate the construct of functional capacity.


Introduction

Several assessment tools have been developed in order to determine patients' disabilities, to assist in claims' decisions, and in return-to-work certifications and strategies. In situations when such an assessment is needed, several factors have to be taken into account, such as individual's proneness to a continued reception of sickness benefit, administration's developed regulations and financial incentives, and the interaction between healthcare providers, employers, and insurers [1,2]. Therefore, to evaluate an

individual's functional status, disability and readiness for work, the use of a standardized tool has been recommended.

One such instrument is a Functional Capacity Evaluation (FCE), which is a clinical instrument using a battery of standardized performance tests to evaluate an individual's ability to safely perform work-related activities [3]. For the purpose of using FCE as a decision-making instrument, its reliability and validity should be established. Although some of the tests display robust evidence for reliability and validity, there is still limited evidence for many

CONTACT Jone Ansuategui Echeita  j.ansuategui.echeita@umcg.nl 

 Supplemental data for this article can be accessed [here](#).

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

protocols and tests [4–6]. A possible reason for this is that functional capacity tests do not only test physical functioning, but they involve personal factors, health status, and environmental factors [7]. The interaction between these factors has been explained through the bio-psycho-social model, where personal or psychological factors, individual health status or biological factors, and environmental or social factors are described [8,9].

Several studies have investigated the potential biopsychosocial factors associated with FCE performance. A Delphi study among scientists, clinicians, and patients examined the most important biopsychosocial factors that were associated with functional capacity results according to the International Classification of Functioning (ICF), Disability and Health framework [10]. Body function, activities, participation, environmental, and personal ICF components were the factors reported to be associated with functional capacity results, but only body function ICF components were part of the main factors. Similarly, two reviews examined evidence on the association of different biopsychosocial factors with performance on functional capacity tests [11,12]. These three studies provided evidence for the association between functional capacity test performance with mainly biological and psychological factors, while social factors were only scarcely described. The social factors were described as, “Factors that make up the physical, social and attitudinal environment in which people live and conduct their lives. These factors are external to individuals and can have a positive or negative influence on the individual’s performance as a member of society, on the individual’s capacity to execute actions or tasks, or on the individual’s body function or structure” [13]. Despite its relevance in patients undergoing FCE testing, it is currently unclear which social factors are associated with FCE performance.

The purpose of the present review was to determine associations between social factors and FCE performance in adults (18–65 years). The findings will provide insights in FCE, which not only will allow clinicians to improve the evaluations and interpretations of the assessment and better design the rehabilitation program, and researchers to further investigate which factors influence FCE performance; but also will add to a better understanding of the construct of functional capacity.

Materials and methods

The present review was performed using best evidence synthesis methods [14,15] and followed the PRISMA reporting guideline [16]. Relevant articles were retrieved from CINAHL, MEDLINE and PsycINFO electronic databases from conception until 31 December 2017. RefWorks (www.refworks.com) was used to remove duplicates and store the retrieved articles.

Systematic search strategy

For FCE terms selection, FCE was split into its composing elements (functional-capacity-evaluation), and synonyms of those elements were searched. Then, the FCE elements, its synonyms and the main FCE capacity tests were combined (i.e., functional capacity, physical performance, and lifting ability). Finally, different types of FCE protocols (i.e., WorkWell, Blankenship, and BTE) were also included in the search terms. The selected terms for social factors were not only related to individuals undergoing assessment; but also to workplace, insurance, legislation, and healthcare environment, and to relationships or interactions developed within those environments. Thus, factors belonging to aspects such as healthcare, sociocultural, socio-demographic, economics and related

policies, social and occupational support, media and technologies were included. No biological or psychological factors such as age, pain, psychological traits or signs, or behaviors of the individual were included in the terms. The search was restricted to peer-reviewed articles and adults (18–65 years). A description of the search strategy can be found in [Supplementary Table S1](#). Furthermore, the reference lists of the eligible full-text articles were screened, and experts in FCE were consulted by email for additional relevant published articles.

Study selection

The selection of the articles was examined by two independent reviewers (JAE and BvH). Of the potential articles, the first 100 in author’s alphabetical order were screened on title by both reviewers and the rest by the principal reviewer (JAE). Of the eligible articles following title screening, the first 50 in author’s alphabetical order were screened on abstract by both reviewers and the remaining by the principal reviewer. When in doubt about the inclusion of an article by title or abstract, it was included and further analyzed. The remaining eligible articles were all full-text analyzed by the two reviewers. The reviewers were blinded to the article authors, publishing journal and keywords in the abstract and full-text screening. In case of disagreement or doubt about the inclusion of an article, consensus was reached by discussion between the reviewers.

Inclusion and exclusion criteria

Eligibility of the studies was examined based on the type of article, population, intervention, comparison, outcomes, and study design. A description of the applied criteria can be found in [Supplementary Table S2](#), and its implementation is represented in [Figure 1](#). Articles were included if all six criteria were applied: (1) were peer-reviewed English, Spanish, French, or Dutch written articles (type of article); (2) reported within- or between-group comparison (comparison); (3) were observational or intervention studies (study design); (4) participants were adults (18–65 years), and were undergoing an FCE test to assess their physical capacity or functional performance (population); (5) the reported outcomes incorporated changes in FCE tests performance (outcome) related to social factors (intervention). There was no date restriction in the search. Articles were excluded if: (1) changes in FCE test performance were not measured with the purpose of assessing participants’ activity, i.e., physical capacity or functional performance; but as a means to assess participation as return-to-work, cognitive traits, or pharmacological/treatment effects (outcome); (2) no social factors were involved in FCE test performance changes (intervention); (3) individuals affected by social factors were not the ones undergoing FCE, such as FCE assessors (population). Books, meetings or conferences, interviews, reviews, master dissertations, letters to the editor, guidelines, and editorials (type of article) were excluded.

Assessment of methodological quality

The methodological quality of each of the studies was independently assessed by two reviewers (the principal author and one of the coauthors), composing a total of six reviewers. The division of the articles among the coauthors was based on the social factors studied. Additionally, having papers assessed by an author of the respective paper was avoided. Disagreements were discussed by the reviewers, until consensus was reached. The methodological quality assessment tool employed for this review was the

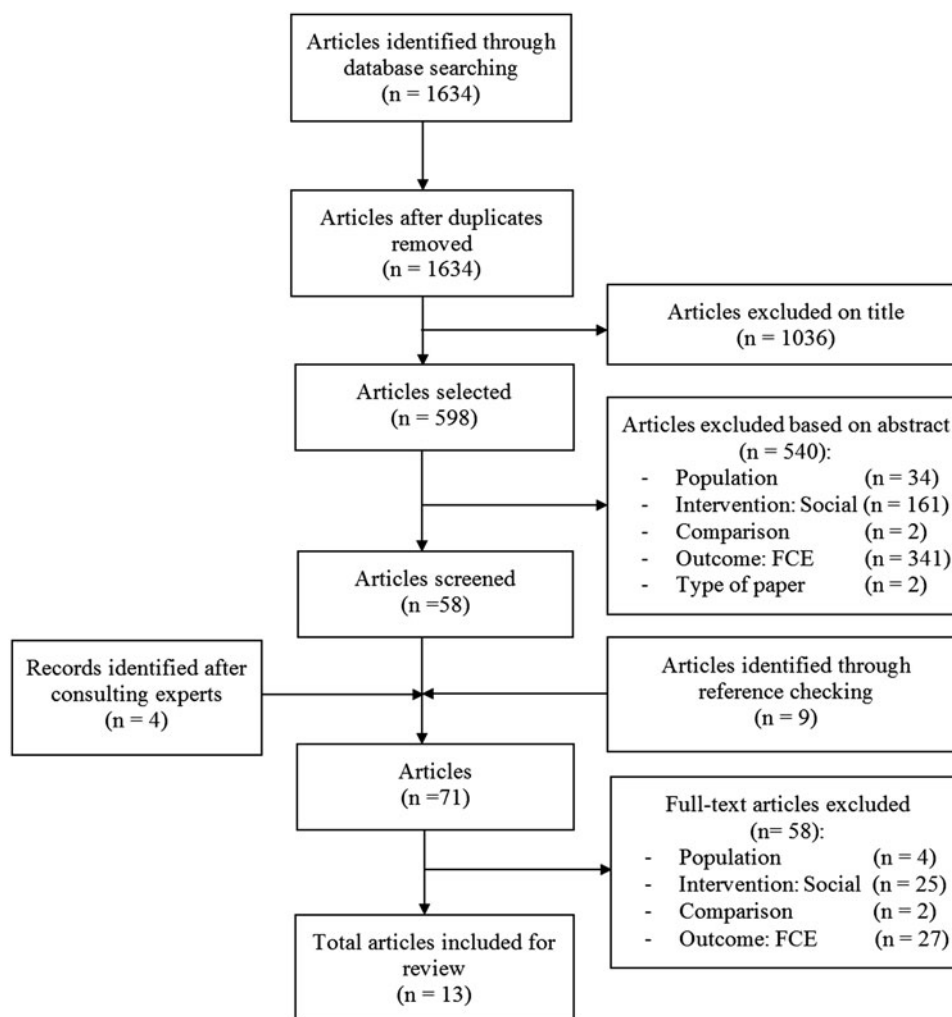


Figure 1. Flow diagram of the study selection procedure.

Table 1. Standard Quality Assessment criteria (QualSyst).

1. Question/objective sufficiently described?
2. Study design evident and appropriate?
3. Method of subject/comparison groups selection or source of information/input variables described and appropriate?
4. Subject (and comparison group, if applicable) characteristics sufficiently described?
5. If interventional and random allocation was possible, was it reported?
6. If interventional and blinding of investigators was possible, was it reported?
7. If interventional and blinding of subjects was possible, was it reported?
8. Outcome and (if applicable) exposure measure(s) well defined and robust to measurement/misclassification bias? Means of assessment reported?
9. Sample size appropriate?
10. Analytic methods described/justified and appropriate?
11. Some estimate of variance is reported for the main results?
12. Controlled for confounding?
13. Results reported in sufficient detail?
14. Conclusions supported by results?

QualSyst (Table 1), for elaboration on the criteria applied: refer to the original article [17]. The QualSyst tool for quantitative studies is a checklist that assesses both the methodology and reporting quality of the articles. It can be used for both cohort and cross-sectional studies without penalization for study design. This checklist consists of 14 items, each of which was scored depending on the degree to which the criterion was met: yes – 2 points, partial – 1 point, no – 0 points, and not applicable (excluded from the calculation of the summary score). The summary score was

obtained by dividing the total score (sum of all the positive items and all the partial items) by the total possible score (28 – (number of N/A × 2)). No further guidance for the classification of the articles was provided; as a result, it was decided that articles were rated high if they scored above the 3rd quartile (75% or over), moderate if they scored between the median and the 3rd quartile (50% and 74%), and low if they scored below the median (49% or less).

Data extraction

From the selected articles, the principal reviewer extracted the information. Accuracy was verified by the reviewer who assessed the methodological quality of the corresponding article. The following details were extracted: study characteristics (design, inclusion/exclusion criteria, type of intervention), sample characteristics (sample size, demographic characteristics, and social variables), and outcome characteristics (type of FCE performed, analyses, and test results). All results were presented in agreement with the original author's judgment on the effect of the social factor on FCE test performance.

Assessment of quality of evidence

The quality of evidence of the outcomes was assessed using an adaptation of the GRADE Working Group guideline (Grading of

Recommendations Assessment, Development and Evaluation) for prognosis research [18] by two reviewers independently. The principal reviewer verified data extraction and graded all the outcomes along with one other reviewer who graded the outcomes of the studies they had already assessed for methodological quality. If necessary, additional information was requested from the corresponding author. In case of disagreement on the grading of an article, consensus was reached through discussion. The level of evidence was graded depending on: phase of investigation, limitations, inconsistency, indirectness, imprecision, publication bias, effect size, and exposure–response gradient. The initial quality of evidence was scored according to the phase of investigation of the study: High (++++) for explanatory studies to understand the prognostic pathways (phase 3 explanatory study) and/or to confirm associations between potential prognostic factor, and the outcome (phase 2 explanatory study); and Moderate (++++) for outcome prediction or explanatory studies to identify associations between potential prognostic factors, and the outcome (phase 1 explanatory study). The initial score was downgraded by one point (−1) or upgraded by one (+1), or stayed the same depending on the risk of bias in the different sections. The overall quality of evidence could be High (++++), Moderate (+++), Low (++), or Very low (+) [18].

Data synthesis

The results of the data extraction were synthesized based on the association of the different social factors with FCE test performance per type of population, either patients or healthy participants. We planned to execute a meta-analyses when two or more studies were retrieved per factor and type of test performed, within a certain comparison. However, this criterion was not met; thus, data synthesis was done descriptively.

Results

Study selection

The results of the literature search and selection are presented in Figure 1. A total of 1634 records were obtained from the search. There were no duplicate records; therefore, 1634 articles were screened on title. This yielded 598 eligible articles. The subsequent abstract screening left 58 articles eligible for full-text assessment. The screening of the reference lists from the eligible full-text articles produced nine articles and consultation of experts in FCE resulted in four additional articles. Full-text assessment left a final selection of 13 articles eligible for quality assessment and data extraction. The articles excluded on abstract and full-text were mainly due to the absence of FCE test and/or social factors. A list with the full-text excluded articles can be found in Supplementary Table S3.

Assessment of methodological quality

The results of the quality assessment for the included articles are displayed in Table 2. Overall, nine articles showed high methodological quality [21–26,29–31], three showed moderate methodological quality [20,27,28], and one showed low methodological quality [19]. The main limitations found were: unclear study design to answer the research question, incomplete description of sampling strategy, reduced reproducible objective criteria of the measurements, sample size appropriateness, insufficient results details, and estimates of the main results. No studies were excluded based on result of methodological quality assessment.

Table 2. Standard Quality Assessment of the eligible articles.

Study	Design	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Quality ^a	
Geisser et al. (2000) [19]	CS	+	−	+/-	+/-	NA	NA	NA	+/-	−	+/-	−	+/-	+	+/-	10/22 (45.5%)	Low
Reneman et al. (2001) [20]	CS	+	+/-	+/-	+	NA	NA	NA	+/-	+/-	+	+/-	NA	+/-	+/-	12/20 (60%)	Moderate
Cutler et al. (2003) [21]	CT	+	+	+	+	NA	NA	NA	+/-	+	+	+	+/-	+/-	+/-	18/22 (81.8%)	High
Gross and Battié (2005) [22]	CS	+	+	+	+	NA	NA	NA	+/-	+	+	+	+	+/-	+	20/22 (90.9%)	High
Reneman et al. (2006) [23]	CS	+	+	+	+	NA	NA	NA	+	+	+	+	+	+	+	22/22 (100%)	High
Kyi et al. (2012) [24]	CS	+	+/-	+/-	+	NA	NA	NA	+	+	+	+	NA	+	+/-	17/20 (85%)	High
Oesch et al. (2012) [25]	CS	+	+	+	+	NA	NA	NA	+/-	+	+	+	+	+	+	21/22 (95.5%)	High
Lakke et al. (2013) [26]	CS	+	+	+/-	+	NA	NA	NA	+	+	+	+	NA	+	+	19/20 (95%)	High
Weir et al. (2013) [27]	CS	+	+/-	+	+/-	NA	NA	NA	+/-	+/-	+/-	+/-	NA	+/-	+	13/20 (65%)	Moderate
Rutherford Owen and Jones Wilkins (2014) [28]	EPF	+	+/-	−	+/-	NA	NA	NA	−	+/-	+	−	NA	+	+	11/20 (55%)	Moderate
Trippolini et al. (2015) [29]	CS	+	+	+	+	NA	NA	NA	+	+	+	+	NA	+	+	20/20 (100%)	High
Lakke et al. (2015) [30]	CRCS	+	+	+/-	+	+	+	+	+/-	+	+	+	+	+	+	26/28 (92.9%)	High
McKay et al. 2017 [31]	CS	+/-	+	+	+	NA	NA	NA	+	+	+	+/-	NA	+	+/-	17/20 (85%)	High

+ : yes; +/- : partial; − : no; NA: not applicable.

CS: cross-sectional; RCS: randomized cross-sectional; CT: cohort; EPF: ex post-factor; CRCS: cluster randomized cross-sectional.

^aQuality: low, < 50%; moderate, 50–74%; high, ≥ 75%.

Study description

The retrieved articles were published between 1 January 2000 and 31 December 2017. Most studies were observational with use of primary data; however, two articles used either stored data from a previous study [23], or selected data from the caseload of a rehabilitation counselor and two attorneys [28]. Participants performing FCE were involved in rehabilitation programs, RTW processes, compensation claims or were healthy workers or students. The mean age of the samples ranged from 36 to 56 years, except two studies in which the examined samples were students (mean age 20.5 and 21.6) [20,30], and one study that collected normative data from participants between 3 and 101 years of age [31]. Information extracted from the studies is summarized in Table 3 for patients and in Table 4 for healthy participants.

Eleven social or environmental factors were identified. Six involved patients: compensation status [19,21], litigation status [19], country of FCE [23], previous job salary [25], type of instructions [27], and primary or mother language [28,29]; three involved healthy participants: environmental conditions [20], time of day (workday) [24], and examiner's fear behavior [30]; and two involved both populations: self-reported workplace social support [22,26] and ethnicity [28,31]. In addition to this, many social characteristics were gathered such as education level, employment situation, and work experience. In most cases, these factors were provided only as sample demographic descriptions; hence, they were not included in the analysis of the studies. FCE measurements were composed of: stoop, climb, crouch, lift, carry, push and pull, forward bent stand, overhead work, grip strength, manual dexterity, ambulation, and sincerity of effort.

Data synthesis

Patients

Compensation status. Two studies analyzed the relation between compensation status and FCE performance [19,21]. One article reported an association between compensated individuals and lower FCE performance (floor-to-waist and waist-to-shoulder lift) [19]; whereas the other article reported that compensation status was associated with a higher rate of failed FCE tests [21]. These associations were maintained for climbing and crouching after adjusting for pain and psychological characteristics. Of the FCE tests the studies measured, only lifting capacity (floor-to-waist lift) was similar in both studies. When the association between compensation status and lifting capacity was analyzed, one of the studies found lower lifting performance, whereas the other article did not find a difference. Overall, FCE performance differed between patients receiving compensation and those who did not. When the relation was further explored it was found that patients receiving compensation had lower performance and more failed FCE tests. However, the only common test in both studies, lifting performance, showed conflicting results between the studies.

Litigation status. Only one study investigated the association between litigation status and FCE performance [19]. Its findings displayed an association between individuals involved in litigation and lower performance of lifting capacity (floor-to-waist and waist-to-shoulder lift) even after accounting for demographic, pain, psychological, and physical characteristics. In general, patients involved in litigation processes had a lower lifting performance.

Self-reported workplace social support. One study addressed this topic and reported no significant association between self-reported workplace social support and FCE performance [22]. The study investigated the relation between the patients' perceptions

of the workplace (measured with Organizational Policies and Practices (OPP) scale) and FCE performance indicators (floor-to-waist lift and number of failed tasks). Patients' perceptions of the workplace were found to have no significant association with performance on the FCE lift test or number of failed tasks. Hence, there was evidence of no relation between self-reported workplace social support and FCE performance. This was consistent with the results found in the study with healthy participants [26].

Country of FCE. The association between country of FCE and FCE performance was studied in one article [23]. This study compared patients FCE tests' results from three different countries (The Netherlands, Canada, and Switzerland), and found differences in lifting and carrying capacity between the Dutch sample and the other two samples. The Dutch sample lifted and carried significantly more than the Canadian and Swiss samples; this association remained significant after controlling for physical and psychological characteristics in lifting capacity. Thus, it appeared that differences in FCE performance between different countries exist.

Previous job salary. One study analyzed the relation between FCE performance and the salary of a patient's previous job [25]. This study reported an association of higher previous job salary with larger walking distance in the 6-min walk test (6MWT), and this association remained significant after correcting for physical and psychosocial variables. The other tests did not show any relationships. In general, a higher salary was related to a larger distance walked, without any other relation to FCE test performances.

Type of instruction. One study investigated the association between different types of instruction for the execution of the 6MWT and its performance [27]. The study reported that the type of instructions given by the examiners had an effect on walked distance. The different instructions were: "walk as far as possible" (standard walk), "walk as fast as you can" (fast walk), "walk at your normal pace" (normal walk), and "walk at a leisurely pace" (leisure walk). Patients walked a mean of 52.7 m more when they were instructed for the fast walk, compared to the standard walk. In general, there was a difference in walked distance when patients were instructed in different ways.

Primary or mother language. Two studies investigated the association of the primary or mother language of the patients on FCE performance [28,29]. Both studies found differences in FCE performance associated with participant's mother language groups in the German canton of Switzerland (German and non-German) and United States (English and non-English). The study comparing primary German and non-German language speakers evidenced that the German-speaking group performed consistently higher on all FCE tests included in the study, i.e., handgrip strength, waist-to-overhead lift, and overhead work and repetitive reaching task [29]. The study comparing primary English and non-English language speakers evidenced that the English-speaking group showed a moderately higher percentage of valid performance effort (70.2%), with respect to the non-English-speaking group (27.3%) [28]. Therefore, evidence consistently showed FCE performance to be higher in patients whose primary or mother languages was the local language.

Ethnicity. One study analyzed the relation between race or ethnicity and FCE sincerity of effort [28]. The study outcomes showed a moderate difference in the number of valid FCE performances between two ethnic groups, White/Caucasian and non-White. The White/Caucasian group showed higher percentage of valid performance effort (72.0%), compared with the non-White group (38.9%). As a result, the authors concluded that differences in validity of FCE performance were related to patient's ethnicity.

Table 3. Results of FCE performance in relation to social factors from patients.

Study	Population	Social Factors	Outcome	Results
Geisser et al. (2000) [19] USA	Chronic disabling back pain assessed at a Spine Program, n = 133	Compensation and litigation status: Compensation Litigation	Progressive Isoinertial Lift Evaluation (PILE) Waist-to-shoulder lift $r = -0.21^{**}$ $\beta = -0.06$ $r = -0.27^{**}$ $\beta = -0.23^{**}$	Compensation and litigation, both showed associations with decreasing floor-to-waist and waist-to-shoulder lifting performances Litigation was a negative predictor of floor-to-waist lift also in the simultaneous regression model.
Cutler et al. (2003) [21] USA	CLBP admitted to a multidisciplinary pain treatment center, n = 188	Compensation status: Worker compensation status	Dictionary of Occupational Titles Residual Functional Capacity battery (DOT-RFC) At admission Climb $\chi^2 = 9.35^{**}$ Crouch $\chi^2 = 13.26^{**}$ Lift $\chi^2 = 5.96$ Stoop $\chi^2 = 6.79^{**}$ Climb $\chi^2 = 2.93$	Workers' compensation status at admission showed differences in the rates of failure of stooping, climbing, and crouching In climbing and crouching logistic regression models, receiving compensation was associated with higher failure rates At discharge, workers' compensation status showed differences in the failure rates of stooping and crouching No apparent relationship between claimant's ratings on OPP scale and FCE performance was found
Gross and Batié (2005) [22] Canada	Low back injured workers for work assessment on claims and RTW decision-making purposes, n = 170	Organizational Policies and Practices (OPP): Claimant's perception of support at workplace Countries	Ismhagen Work System Functional Capacity Evaluation (IWS-FCE) Number of failed tasks ^a 9 out of 25 (5.7) $r = 0.03-0.08$	Dutch patients showed higher FCE lifting and carrying performance than Canadian and Swiss patients. In final lifting regression models, this relationship remained significant
Reneman et al. (2006) [23] The Netherlands, Canada, and Switzerland	CLBP patients: - In The Netherlands for an admission assessment. - In Canada for workers' compensation claimants. - In Switzerland for pre-treatment assessment. n = 564	The Netherlands (n = 121) Canada (n = 273) Switzerland (n = 170)	Ismhagen Work System Functional Capacity Evaluation (IWS-FCE) Horizontal lift (kg) ^a Waist-to-overhead lift (kg) ^a Right-hand carry (kg) ^a Left-hand carry (kg) ^a Front carry (kg) ^a	
Oesch et al. (2012) [25] Switzerland	Patients with chronic NSLBP referred for fitness-for-work evaluation on three rehabilitation centers. n = 126	Previous job salary (CHF)	Ismhagen Work System Functional Capacity Evaluation (IWS-FCE) Forward bent standing (s) Grip strength dominant hand (kg) B = 0.00 (0.00)	In the regression model, previous job salary was associated with larger walking distance. This relation was also significant in the multiple regression models. No other relations were found
Weir et al. (2013) [27] USA	Patients recruited from an Advanced Lung Disease Clinic, n = 24	Alternative instruction: Standard walk Fast walk Normal walk Leisure walk	6-min walk test (m) ^b 494 (331-679) 547 (417-779)** 354 (274-428) 413 (291-528)	Greater distance was walked during the fast walk compared with the standard walk
Rutherford Owen and Jones Wilkins (2014) [28] USA	Cases from a rehabilitation counselor and two attorneys, regularly involved in the workers compensation system, n = 68	Ethnicity and Primary language: White/Caucasian vs. non-White (n = 50 vs. n = 18) English vs. non-English (n = 57 vs. n = 11)	Sincerity of effort measures for FCE Validity of FCE performances $\chi^2_{(1, 67)} = 62^{**}$ ϕ (CI): -0.30 (-0.54, -0.04) $\chi^2_{(1, 67)} = 7.33^{**}$ ϕ (CI): -0.33 (-0.5, -0.09)	There were differences in sincerity of effort scores between ethnic and primary language groups. White/Caucasian workers and English speakers showed a higher percentage of valid performance effort. The relation between ethnicity and primary language with sincerity of effort scores showed medium effect sizes
Trippolini et al. (2015) [29] Switzerland	WAD grade I or II patients for rehabilitation assessment at the rehabilitation clinic, n = 314	Primary language: German (n = 152) Non-German (n = 162)	WorkWell Functional Capacity Evaluation (IWS-FCE) Handgrip strength (kgF) ^a Waist-to-overhead lift (kg) ^a Overhead work (s) ^a Repetitive task (s) ^a	Relationship between FCE performance and patient's primary language spoken was found. The German-speaking group showed higher performance in all four FCE tests

LBP: low back pain; CLBP: chronic low back pain; NSLBP: non-specific low back pain; WAD: Whiplash associated disorders.

Significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

^aMean (SD). ^bMean (range). ^cCultural background was measured based on the primary language, i.e., the mother language.

ϕ : the phi coefficient is a Pearson's product-moment coefficient of correlation calculated on two nominal-dichotomous variables when the categories of both variables are coded 0 and 1. This measure is used to estimate the magnitude of association in 2×2 contingency tables, as in the chi-squared test ($\phi = \sqrt{\frac{\chi^2}{n}}$).

Table 4. Results of FCE performance in relation to social factors from healthy participants.

Study	Population	Social Factors	Outcome				Results
			Isernhagen Work System	Functional Capacity Evaluation (IWS-FCE) – Production rate			
Reneman et al. (2001) [20] The Netherlands	Healthy students, $n = 24$	Environmental conditions Normal Max. production Noise	Overhead work (n/min) ^a	Kneeling (n/min) ^a	Production rate	The three environmental conditions did not appear to influence FCE performance	
			6.6 (1.3)	6.4 (1.0)			
			7.3 (1.9)	6.6 (1.2)			
			7.0 (2.5)	6.4 (1.1)			
Kyi et al. (2012) [24] Canada	Healthy working age adults, $n = 50$	Workday Morning Afternoon	FCE			Participant's performance improved after a day's work; however, this was only significant on two of FCE tests. The effect of time of day difference on manual dexterity was large, whereas on walk speed was very small	
			Right grip (kg) ^a	Left grip (kg) ^a	Manual dexterity (s) ^a		50-foot walk (s) ^a
			47.33 (13.7)	45.58 (13.7)	50.92 (5.6)*	7.02 (0.85)***	
			48.43 (14.8)	46.11 (14.5)	45.92 (5.0)	6.83 (1.0)	
Lakke et al. (2013) [26] The Netherlands	Healthy workers, $n = 403$	Perception of workplace support (QAEW) Co-worker support Supervisory Support	FCE			No relations were found between FCE performances and workplace social support	
			Low lift (kg)	Overhead lift (kg)	Carry (kg)		Forward bent standing (s)
			$\rho = -0.03$	$\rho = 0.01$	$\rho = -0.05$	$\rho = 0.00$	$\rho = -0.08$
			$\rho = 0.02$	$\rho = 0.04$	$\rho = 0.05$	$\rho = -0.01$	$\rho = -0.07$
Lakke et al. (2015) [30] The Netherlands	First- and second-year physical therapy healthy students, $n = 256$ participants $n = 24$ examiners	Examiner's fear behavior Low fear ($n = 132$) High fear ($n = 124$)	WorkWell FCE			Participants with the low-fear examiner scored higher than participants with the high-fear examiner. In the final model, this association remained significant	
			Lifting capacity (kg) ^a				Functional Dexterity Test (s) ^a
			39.6 (16.4)***				
			B = 14.41 (3.6)**				
			32.1 (13.6)				
McKay et al. (2017) [31] Australia	Healthy population, $n = 399$	Ethnicity British/European ($n = 265$) non-British/ European ($n = 134$)	WorkWell FCE			Differences in performance between the ethnic groups were found in one of the FCE tests, British/Europeans walked significantly longer distance. No differences were found for the other FCE tests	
			6-Min Walk Test (m) ^a	9-Hole Peg Test (s) ^a			
			711.7 (91.4)*	17.7 (2.5)	20.5 (4.2)		
			693.1 (80.6)	17.5 (2.4)	20.6 (4.4)		

QAEW, Dutch Questionnaire on the Appreciation and Evaluation of Work.

Significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.^aMean (SD). ^bMean (range).

Table 5. Quality of evidence of social factors addressed by eligible articles (GRADE).

Outcome	Phase of investigation (RCT, Phase3 or 2/ Phase1)	Initial score	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Large effect	Exposure-response	Overall quality
Compensation status	2 (0/2)	+++	-1	No serious inconsistency	No serious indirectness	-1	Undetected	No large effect	No exposure-response gradient	+
Litigation status	1 (0/1)	+++	-1	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	No large effect	No exposure-response gradient	++
Country of FCE	1 (0/1)	+++	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	+1	No exposure-response gradient	++++
Previous salary	1 (0/1)	+++	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	No large effect	No exposure-response gradient	+++
Alternative instruction	1 (0/1)	+++	-1	No serious inconsistency	No serious indirectness	-1	Undetected	+1	No exposure-response gradient	++
Primary or mother language	2 (0/2)	+++	-1	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	Results pointing in the same direction	No exposure-response gradient	+
Ethnicity	2 (0/2)	+++	-1	-1	-1	No serious imprecision	Undetected	No large effect	No exposure-response gradient	0
Workplace social support	2 (0/2)	+++	-1	No serious inconsistency	-1	No serious imprecision	Undetected	+1	No exposure-response gradient	++
Environmental conditions	1 (0/1)	+++	-1	No serious inconsistency	-1	No serious imprecision	Undetected	No large effect	No exposure-response gradient	+
Time of day	1 (0/1)	+++	-1	No serious inconsistency	-1	-1	Undetected	No large effect	No exposure-response gradient	+
Examiner's behaviors	1 (1/0)	+++	No serious limitations	No serious inconsistency	-1	No serious imprecision	Undetected	+1	No exposure-response gradient	++++

RCT: randomized controlled trial.

++++, high-quality evidence; +++, moderate-quality evidence; ++, low-quality evidence; + or 0, very low-quality evidence; -1, quality of evidence is downgraded by one +; +1, quality of evidence is upgraded by one +.

Healthy participants

Environmental conditions. One study analyzed the association between the environmental conditions and FCE performance [20]. Environmental conditions were operationalized as: normal (tests performed in consonance with the protocol), maximal production (tests of screws/bolts manipulation performed as fast as possible), and noise (tests performed with annoying volume level). The study results displayed no differences in FCE performance between three environmental conditions. Therefore, the evidence showed there was no relation between these environmental conditions and FCE performance.

Time of day (workday). One study investigated the association of time of day or effect of workday, and FCE performance [24]. FCE measurements were taken in the morning and later in the afternoon after a day's work. Participants were able to perform relevantly faster on the manual dexterity test (5 s), and to walk faster on the 50-foot walk test (0.19 s) in the afternoon. Opposed to this, handgrip performances did not show differences. In general, the afternoon (after work) measurements were significantly associated with an improvement in FCE performance, specifically on manual dexterity and walking velocity; no other FCE test performance showed a significant association with this factor.

Self-reported workplace social support. One study addressed the topic of self-reported workplace social support and the results reported no associations between self-reported workplace social support and FCE performance [26]. This study examined the relation between healthy participants' self-reported workplace social support: coworker and supervisory support (measured with the Dutch Questionnaire on the Appreciation and Evaluation of Work (QAEW)), and FCE performance. Participant's coworker and supervisory support was not significantly associated with performance on any of the measured FCE tests. Hence, it was evidenced there was no relation between self-reported workplace social support and FCE performance. This was consistent with the results found in the study with patients [22].

Examiner's fear behavior. Fear behavior issue was evaluated in one study [30]. The study reported that examiners' fear behavior during testing significantly influenced lifting performance of the participants. High fear of injury examiners showed a greater guarding behavior by expressing more frequently symptom-focused talks, lifting avoidance, reassurance, ergonomic verbal instructions, procedural talk, and examiner's decisions. Participants evaluated by a low-fear examiner lifted 7.4 kg more than participants with a high-fear examiner; the association remained after correcting for participant's physical and personal characteristics. Overall, there was a difference on lifting performance between the groups tested with high- or low-fear examiners.

Ethnicity. One study investigated the association of ethnicity and FCE performance [31]. The study compared a British/European ethnic group with non-British/Europeans in several tests. Of the tests that belong to the FCE battery, differences between the groups were only observed on the 6MWT. The British/European ethnic group (711.7 m) performed significantly better than the non-British/European (693.1 m). In general, the British/European ethnic group walked a longer distance in the 6MWT. No other FCE test performance showed a significant difference between ethnic groups.

Assessment of quality of evidence

The quality of evidence was graded for 11 factors: compensation status, litigation status, self-reported workplace social support, country of FCE, previous job salary, type of instruction, primary or

mother language, ethnicity, environmental conditions, time of day (workday), and examiner's fear behavior. The results of the quality of evidence assessment for the included articles are displayed in Table 5. Overall, the level of evidence for the factors studied in patients was high for country of FCE [23], moderate for previous job salary [25], low for litigation status [19] and type of instruction [27], and very low for compensation status [19,21] and primary or mother language [28,29]. In healthy participants, the level of evidence for the factors studied was high for examiner's fear behavior [30], and very low for time of day (workday) [24] and environmental conditions [20]. Only two factors, self-reported workplace social support [22,26] and ethnicity [28,31], were studied in both patients and healthy participants, and showed low and very low level of evidence respectively. The initial score was moderate for 10 factors due to being cross-sectional observational studies, and the phase of investigation they belonged to was 1. The subtraction of + from was a consequence of: limitations, inconsistency, indirectness, and imprecision. The addition of + was due a large magnitude of effect.

Discussion

The main purpose of this study was to systematically review and summarize the association between various social factors and FCE performance in adults. Thirteen studies were found to investigate the association between FCE performance and 11 different social factors. There was high quality of evidence that patients performing FCE tests in different countries [23] had significant differences in FCE results, and healthy participants tested in presence of a low-fear examiner [30] lifted significantly greater weight in the floor-to-waist lift test. There was moderate quality of evidence that higher previous job salary was significantly associated with larger distance walked in the 6MWT [25]. There was low or very low quality of evidence that patients not involved in compensation [19,21] and/or litigation [19] processes, and healthy participants tested after a day's work [24] performed significantly better on FCE tests; also, patients with differences in the instruction [27] and in their primary or mother language [28,29], and both patients and healthy participants with different ethnicity [28,31], had significantly different FCE performances. There was an absence of association of workplace social support reported by patients and healthy participants [22,26] and the different environmental conditions, in which FCE was undertaken [20], with FCE performance.

Although social characteristics such as employment, education, claims status, and disability have served as sample characteristics in FCE research, only two reviews have analyzed the association between psychosocial factors and FCE [11,12]. Both reviews found scarce evidence on the association of compensation, litigation, and secondary gain with FCE performance. To our knowledge, this is the first systematic review that specifically focuses on social factors belonging to the individuals' legislative and insurance, workplace, healthcare, and personal or cultural systems, thereby providing a broader overview of a large variation of social factors associated with FCE performance. This is consistent with Loisel's arena of work disability, which emphasizes the importance of these domains and multiple factors [32].

In agreement with the definition of FCE, FCE results should be interpreted considering patients' physical functioning, personal factors, health status, and environmental factors [7]. The findings of this review contribute to our understanding of associations between social factors and FCE performance. FCE results must be interpreted in a broader context to provide a more comprehensive representation of patients' performance. As a consequence of

differences between jurisdictions, reference values may not be generalizable. Country specific reference values may need to be developed, which may lead to better interpretations of FCE results for each patient. As a result, optimized recommendations based on these as well as more effective rehabilitation programs due to a better targeting of patients may be possible. Additionally, it is important to outline that some associations between FCE and social factors may be adaptable. For example, clinicians assessing FCE, being a social factor themselves, could adapt their attitudes or behaviors, or instructions which may influence patients' performance.

There are some limitations in this systematic review that need to be considered. To begin with, even though a broad search strategy was applied, we cannot be fully confident that relevant articles have been missed in the search. However, it is assumed that the number of absent articles is minimal and that they would not likely alter the conclusions. Another limitation is the broadness of social and environmental factors and their inherent relationship with psychosocial factors. As some social factors produce a psychological effect in the individual, it is difficult to discern the boundaries between these domains. For example, while fear avoidance could be considered a psychological factor, this may (partly) be induced by social factors such as healthcare professional behaviors [33]. Therefore, there might be some social-psychosocial factors not included in the review; nevertheless, we trust that the main and more obvious social factors have been included. Additionally, the methodological quality assessment tool used in this review was limited due to the fact that it does not define the benchmarks between high-, moderate-, and low-quality studies [17]. For that purpose, we established cutoff values at the median and 3rd quartile, a choice that has been previously implemented in similar situations [34]. Many different FCE tests and protocols were considered in this systematic review, which have varying evidence of reliability and validity. As a result, variability across protocols may have influenced the trustworthiness of our results. A final limitation is the difficulty inherent in measuring important social factors such as patient's ethnicity. One study analyzed the relation of patient's ethnicity to "valid" FCE performances instead of to FCE performance itself [28]. However, there was no definition or description of what constituted a "valid" FCE performance. Secondly, while the authors concluded that differences in validity of FCE performance were related to patient's ethnicity, they also stated: "Examination of actual administrations across ethnic and language groups was not undertaken in this study and may provide information regarding causes of invalid assessments for non-English speaking workers. In this sample, it is unknown which assessments were conducted with an English/Spanish translator" [28, p. 59]. Given these limitations we concluded that there was no convincing evidence that ethnicity by itself is associated with FCE performance.

There are also strengths of this study to be accounted for. The procedure has been transparent, following a systematic search strategy in various databases, using reliable methodological and quality appraisal assessment tools, and best evidence synthesis methods.

Quality of evidence is viewed as of paramount importance for the interpretation of the study results. This was high for associations with the country where FCE is performed. Neither physical nor psychological characteristics of the investigated population explained the performance difference in this study [23]. It can therefore be hypothesized that certain characteristics of these countries, such as health care, personal or cultural systems may explain the performance difference. High quality of evidence was also found for an association of examiners fear behavior with FCE

performance which can be also viewed as an aspect of the health care, personal or cultural systems. Moderate quality of evidence was found for salary in the previous job which is viewed as one aspect of workplace system factors. All other findings of associations of legislative and insurance (compensation and litigation) systems as well as of health care (type of instruction), personal or cultural (primary or mother language and ethnicity), and workplace (workplace social support, environmental condition, and time of day) system factors with FCE performance were of low or very low quality.

Overall, the findings showed that social factors have a bearing in FCE performance and these should be carefully considered in future studies. First, the FCE tests as well as the social factors examined were not uniformly assessed in all studies: the majority of the included articles studied a lifting capacity test but only 4 out of the 11 social factors were considered in more than one study. Second, 5 studies did not measure the typical FCE test target population, instead healthy workers', healthy population, or students' populations were assessed [20,24,26,30,31]. Although the quality of evidence in the factors they studied was downgraded for this, the generalizations to patients with musculoskeletal diseases should be made with caution. Finally, the studies included were mainly cross-sectional in its majority and, therefore, any causal relations should be avoided. Future research should further investigate on which and how social factors, also within a biopsychosocial context, influence FCE. Because social factors serve as a description of certain characteristics of the participants undergoing FCE tests, guidance to further investigate FCE construct could also be obtained. Based on the quality of evidence results from our review, special attention should be paid to personal or cultural, health care, and workplace systems (as represented by country of FCE, clinician's fear beliefs, and previous job salary).

Conclusions

A variety of social factors were found to be associated with performance during FCEs. The considerable heterogeneity found in the measures, the populations, and the methods, prevent robust conclusions about independent associations of social factors on FCE performance. Further research on social factors is required to have a more comprehensive understanding of FCE performance.

Acknowledgements

The authors would like to express their gratitude to A. R. den Otter for his assistance and recommendations to the development of the review.

Disclosure statement

Five of the review coauthors (D.P. Gross, J. Kool, P. Oesch, M.A. Trippolini, and M.F. Reneman) were also authors and/or coauthors in some of the included articles. According to the Cochrane Review Guidelines and to avoid conflict of interest these coauthors were not involved in either the methodological quality or the quality of evidence appraisals of their studies, nor in the data analysis of this study [35]. All authors declare they have no conflict of interest.

Funding

This systematic review received no funds for its development.

References

- [1] Leijon O, Josephson M, Österlund N. Sick-listing adherence: a register study of 1.4 million episodes of sickness benefit 2010–2013 in Sweden. *BMC Public Health*. 2015;15:380.
- [2] Franche RL, Corbière M, Lee H, et al. The Readiness for Return-To-Work (RRTW) scale: development and validation of a self-report staging scale in lost-time claimants with musculoskeletal disorders. *J Occup Rehabil*. 2007;17:450–472.
- [3] Isernhagen SJ. Functional capacity evaluation: rationale, procedure, utility of the kinesio-physical approach. *J Occup Rehab*. 1992;2:157–168.
- [4] Innes E. Reliability and validity of functional capacity evaluations: an update. *Int J Disabil Manag*. 2006;1:135–148.
- [5] Lechner DE, Page JJ, Sheffield G. Predictive validity of a functional capacity evaluation: the physical work performance evaluation. *Work*. 2008;31:21–25.
- [6] Ting W, Wessel J, Brintnell S, et al. Validity of the Baltimore Therapeutic Equipment Work Simulator in the measurement of lifting endurance in healthy men. *Am J Occup Ther*. 2001;55:184–190.
- [7] Soer R, van der Schans CP, Groothoff JW, et al. Towards consensus in operational definitions in functional capacity evaluation: a Delphi survey. *J Occup Rehabil*. 2008;18:389–400.
- [8] Engel GL. The need for a new medical model: a challenge for biomedicine. *Holis Med*. 1989;4:37–53.
- [9] Waddell G. Preventing incapacity in people with musculoskeletal disorders. *Br Med Bull*. 2006;77:55–69.
- [10] Lakke SE, Wittink H, Geertzen JH, et al. Factors that affect functional capacity in patients with musculoskeletal pain: a Delphi study among scientists, clinicians, and patients. *Arch Phys Med Rehabil*. 2012;93:446–457.
- [11] Geisser ME, Robinson ME, Miller QL, et al. Psychosocial factors and functional capacity evaluation among persons with chronic pain. *J Occup Rehabil*. 2003;13:259–276.
- [12] van Abbema R, Lakke SE, Reneman MF, et al. Factors associated with functional capacity test results in patients with non-specific chronic low back pain: a systematic review. *J Occup Rehabil*. 2011;21:455–473.
- [13] World Health Organization (WHO). ICF: International Classification of Functioning Disability and Health. Geneva: WHO; 2002.
- [14] Slavin RE. Best evidence synthesis: an alternative to meta-analytic and traditional reviews. *Educ Res*. 1986;15:5–11.
- [15] Slavin RE. Best evidence synthesis: an intelligent alternative to meta-analysis. *J Clin Epidemiol*. 1995;48:9–18.
- [16] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264–269.
- [17] Kmet LM, Lee RC, Cook LS. Standard Quality Assessment criteria for evaluating primary research papers from a variety of fields. Edmonton (AB): Alberta Heritage Foundation for Medical Research AHFMR (CA); 2004.
- [18] Huguet A, Hayden JA, Stinson J, et al. Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Syst Rev*. 2013;2:71.
- [19] Geisser ME, Haig AJ, Theisen ME. Activity avoidance and function in persons with chronic back pain. *J Occup Rehabil*. 2000;10:215–227.
- [20] Reneman MF, Joling CI, Soer EL, et al. Functional capacity evaluation: ecological validity of three static endurance tests. *Work*. 2001;16:227–234.
- [21] Cutler RB, Fishbain DA, Steele-Rosomoff R, et al. Relationships between functional capacity measures and baseline psychological measures in chronic pain patients. *J Occup Rehabil*. 2003;13:249–258.
- [22] Gross DP, Battie MC. Factors influencing results of functional capacity evaluations in workers' compensation claimants with low back pain. *Phys Ther*. 2005;85:315–322.
- [23] Reneman MF, Kool J, Oesch P, et al. Material handling performance of patients with chronic low back pain during functional capacity evaluation: a comparison between three countries. *Disabil Rehabil*. 2006;28:1143–1149.
- [24] Kyi MT, Fischer JA, Penner J, et al. Physical effect of work on healthy individuals: implications for FCE testing. *Work*. 2012;42:233–239.
- [25] Oesch P, Meyer K, Jansen B, et al. What is the role of “nonorganic somatic components” in functional capacity evaluations in patients with chronic nonspecific low back pain undergoing fitness for work evaluation? *Spine*. 2012;37:E243–E250.
- [26] Lakke SE, Soer R, Geertzen JHB, et al. Construct validity of functional capacity tests in healthy workers. *BMC Musculoskelet Disord*. 2013;14:13.
- [27] Weir NA, Brown AW, Shlobin OA, et al. The influence of alternative instruction on 6-min walk test distance. *CHEST J*. 2013;144:1900–1905.
- [28] Rutherford OT, Jones WM. Sincerity of effort differences in functional capacity evaluations. *J Rehabil*. 2014;80:53–61.
- [29] Trippolini MA, Dijkstra PU, Geertzen JHB, et al. Construct validity of functional capacity evaluation in patients with Whiplash-associated disorders. *J Occup Rehabil*. 2015; 25:389–400.
- [30] Lakke SE, Soer R, Krijnen WP, et al. Influence of physical therapists' kinesiophobic beliefs on lifting capacity in healthy adults. *Phys Ther*. 2015;95:1224–1233.
- [31] McKay MJ, Baldwin JN, Ferreira P, et al. Reference values for developing responsive functional outcome measures across the lifespan. *Neurology*. 2017;88:1512–1519.
- [32] Loisel P, Buchbinder R, Hazard R, et al. Prevention of work disability due to musculoskeletal disorders: the challenge of implementing evidence. *J Occup Rehabil*. 2005;15:507–524.
- [33] Darlow B, Fullen BM, Dean S, et al. The association between health care professional attitudes and beliefs and the attitudes and beliefs, clinical management, and outcomes of patients with low back pain: a systematic review. *EJP*. 2012;16:3–17.
- [34] Brouwers MC, Johnston ME, Charette ML, et al. Evaluating the role of quality assessment of primary studies in systematic reviews of cancer practice guidelines. *BMC Med Res Methodol*. 2005;5:8.
- [35] Furlan AD, Pennick V, Bombardier C, et al. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine*. 2009;34:1929–1941.