



**Bias caused by reliance on patient-reported outcome measures in non-blinded randomized trials: an in-depth look at exercise therapy for chronic fatigue syndrome.**

Journal:	<i>Fatigue: Biomedicine, Health &amp; Behavior</i>
Manuscript ID	RFTG-2020-0028
Manuscript Type:	Regular Article
Date Submitted by the Author:	06-Nov-2020
Complete List of Authors:	Tack, Michiel; None, Tuller, David; University of California Berkeley School of Public Health Struthers, Caroline; University of Oxford Centre for Statistics in Medicine
Keywords:	chronic fatigue syndrome, graded exercise therapy, patient-reported outcome measures, bias, blinding

SCHOLARONE™  
Manuscripts

**Title page**

**Title**

Bias caused by reliance on patient-reported outcome measures in non-blinded randomized trials: an in-depth look at exercise therapy for chronic -fatigue syndrome.  
~~Bias caused by reliance on patient-reported outcome measures in non-blinded randomized trials: exercise therapy for chronic fatigue syndrome as a case study.~~

**Authors**

Michiel Tack

*ME/CFS patient not affiliated with an academic institution. Sint-Laurentiusstraat 87, 9700 Oudenaarde Belgium.*

ORCID ID: <https://orcid.org/0000-0002-5249-6963>

Corresponding author

Email: [tackmichiel@gmail.com](mailto:tackmichiel@gmail.com)

Telephone: 0494621114

David M. Tuller

*Center for Global Public Health, School of Public Health, University of California, Berkeley; Berkeley, California, United States of America.*

Contact: [davetuller@berkeley.edu](mailto:davetuller@berkeley.edu)

Caroline Struthers

*Centre for Statistics in Medicine; Nuffield Dept of Orthopaedics Rheumatology and Musculoskeletal Sciences, University of Oxford: Oxford, Oxfordshire, United Kingdom of Great Britain and Northern Ireland.*

ORCID ID: <https://orcid.org/0000-0003-0848-9894>

Contact: [caroline.struthers@csm.ox.ac.uk](mailto:caroline.struthers@csm.ox.ac.uk)

For Peer Review Only

**Bias caused by reliance on patient-reported outcome measures in non-blinded randomized trials: exercise therapy for chronic fatigue syndrome as a case study**

**Background:** Several randomized trials have reported that graded exercise therapy (GET) is an effective treatment for chronic fatigue syndrome (CFS). These trials were not blinded and relied on patient-reported outcome measures (PROMs). We investigate whether bias introduced by this study design influenced the results.

**Methods:** We extracted standardized mean differences from the most recent meta-analysis on exercise therapy for CFS to analyse their size, consistency over time, and congruence with objective measurements. A narrative review methodology was used to examine mediation analyses, plausible mechanisms of improvement, and risk of response bias.

**Results:** Patient-reported improvements in exercise trials for CFS tend to be small, transient, and poorly supported by objective measurements. The risk of expectancy effects and response bias was high as patients were actively encouraged to adopt a positive attitude towards exercise therapy. Mediation analyses suggest that self-reported improvements in fatigue and physical function are not mediated by objective measures of fitness.

**Conclusions:** Treatment effects seen in exercise trials for CFS could be the result of bias associated with the use of PROMs in non-blinded trials. This might explain the discrepancy between positive results reported in randomized trials and views on exercise therapy expressed by patient organizations. We hope that this case study furthers critical assessment of patient-reported improvements in areas of medicine where blinding of therapists and trial participants faces practical limitations.

**Keywords:** chronic fatigue syndrome, graded exercise therapy, patient-reported outcome measures, minimal clinically important difference, bias, blinding.

**Background**

Chronic fatigue syndrome (CFS) is a chronic, debilitating illness affecting an estimated

0.2 [1] to 0.4 [2] percent of the population. Patients suffer from persistent fatigue, cognitive impairment, post-exertional malaise and a range of other symptoms. [3,4] The pathology of CFS is poorly understood [5] and few treatment options are available to patients and clinicians. [6] Other names, such as myalgic encephalomyelitis (ME) [7] or systemic exertion intolerance disease (SEID) [4], have been used to refer to this patient population. We use the term CFS to be consistent with the studies we discuss in this paper.

Several randomized trials have reported that graded exercise therapy (GET) is an effective treatment for patients with CFS. [8,9] GET is designed to help patients find a sustainable baseline of physical activity from which a gradual increase in exercise is undertaken with the help of a healthcare professional. [10] Patient organizations, however, have raised concerns about the use of GET. [11] In multiple surveys, CFS patients report that they found GET unhelpful or even harmful. [12,13] This discrepancy between the results of randomized trials and the concerns expressed by patient organizations requires explanation.

Randomized trials of GET in CFS focused on patient-reported outcome measures (PROMs) such as fatigue questionnaires when neither patients nor therapists could be blinded to treatment allocation. Both Edwards J [14] and Wilshire et al. [15] have argued that this trial design introduces bias that could distort the main outcomes. Patients who know they are receiving an active intervention rather than a passive control such as “treatment as usual” might be more optimistic about its effect on their health (expectancy bias), or report symptoms according to what they think will please the investigators (response bias). [16,17] Following a 12-week course of exercise therapy involving weekly appointments with a health professional, patients might also be unwilling to acknowledge a lack of improvement due to invested time and energy

(sunk-cost fallacy). [16] In this study, we investigate whether bias associated with the use of PROMs in unblinded trials could account for the improvements reported in exercise trials for CFS.

**Methods**

We used the most recent meta-analyses, reported in the Cochrane review on exercise therapy for CFS (version 8, 2019), to extract an independent calculation of mean differences and standardized mean differences. The Cochrane review includes eight randomized trials of GET with data from 1518 participants. [18] All statements in this study refer to the main comparison of exercise therapy versus a passive control condition (treatment as usual or relaxation/flexibility therapy).

To evaluate their size, patient-reported improvements in exercise trials for CFS are compared with estimates of the minimal clinically important difference (MCID), defined as “the smallest difference in a measurable clinical parameter that indicates a meaningful change...”[19] A systematic search of PubMed is used to obtain MCID estimates for the main outcome measures used in the eight exercise trials for CFS (Supplementary Material).

Our analysis focuses on the effect of not blinding patients and healthcare providers. The blinding status of other persons involved in the trial such as outcome assessors or statisticians is outside the scope of this study.

**Results**

***1. Post-treatment differences are comparable to MCID estimates***

A 2014 review by Hróbjartsson et al. on randomized trials that compared blinded and non-blinded groups, found that the average difference in effect size for PROMs was 0.56. [20] In groups where patients weren't blinded, the reported effect sizes were

inflated by approximately half a standard deviation. This is similar to the size of standardized mean differences reported in the Cochrane review (range: 0.17-0.66). [18] It is however difficult to compare blinded and non-blinded effect sizes of behavioral interventions and the review by Hróbjartsson and colleagues focused on interventions such as acupuncture that may have limited comparability with GET.

The post-treatment effect size for fatigue in the Cochrane review corresponds to a 3.4 point reduction when re-expressed on the 33-point Chalder Fatigue Scale. [18] This is within the range of reported MCID estimates for this scale (range 1.4-4, table 1). In CFS studies, MCIDs of 2 [8] and 3 [21] points have been used. Ridsdale and colleagues used an MCID of 4 points in patients with chronic fatigue, explaining that “our consensus view was that a difference of less than four, using a Likert scale, is not important.”[22] The creators of the Chalder Fatigue Scale were included in this consensus view.

**Table 1. Estimates of the minimal clinically important difference (MCID) for the Chalder Fatigue Scale (score 0-33).**

Study	Population	Sample size	Method	MCID
Ridsdale 2001 [22]	Chronic fatigue	N/A	Consensus-based	4
Goligher 2008 [23]	Systemic lupus erythematosus	n = 80	Anchor-based	2.3
Pouchot 2008 [24]	Rheumatoid arthritis	n = 61	Anchor-based	3.3
Petterson 2015 [25]	Systemic lupus erythematosus	n = 51	Anchor-based	1.4
Nøstdahl 2018 [26]	Postsurgery fatigue	n = 442	Anchor-based	3.4

For physical function, measured by a subscale of the 36-Item Short-Form Health Survey (SF-36, score 0-100) [27], the mean difference across trials is 13.10 points.

Considerable heterogeneity was found, driven largely by one trial (Powell et al. 2001), an outlier that reported a “remarkably positive result.” [18] A sensitivity analysis excluding this trial reduced the mean post-treatment difference between the exercise and passive control group to 7.37 points. [18] This is below most MCID estimates for this scale. In adult CFS patients an MCID of 8 points has been used [8,28], while Bridgen et al. estimated the MCID for adolescent CFS patients to be 10 points. [29] Similar MCIDs have been proposed in other chronic illnesses, although estimates vary widely (range 0.76-22.1, table 2).

**Table 2. Estimates of the minimal clinically important difference (MCID) for the physical functioning subscale of the 36-Item Short Form Health Survey (score 0-100).**

Study	Population	Sample size	Method	MCID
Kosinski 2000 [30]	Rheumatoid arthritis	n = 693	Anchor-based	6.4-8.4
Ward 2014 [31]	Rheumatoid arthritis	n = 243	Anchor-based	±16.9*
Hays 2015 [32]	Rheumatoid arthritis	n = 451	Anchor-based	±0.8-9.2*
de Groot 2006 [33]	Multiple sclerosis	n = 156	Anchor-based	8.5-8.6
Robinson 2009 [34]	Relapsing-remitting multiple sclerosis	n = 249	Anchor-based	±10-22.1*
Devilliers 2015 [35]	Systemic lupus erythematosus	n= 185	Anchor-based	1.9 – 4.9
McElhone 2016 [36]	Systemic lupus erythematosus	n = 101	Anchor-based and distribution- based	2.4-13.4
Coteur 2009 [37]	Crohn’s disease	n = 1328	Anchor-based method	8.3-8.7
Erez 2016 [38]	Chronic kidney disease	n = 73	Anchor-based and distribution- based	7.4-7.9



Wyrwich 2007 A [39]	Chronic Obstructive Pulmonary Disease	n = 610	Anchor-based, distribution- based and consensus-based	1-10
Wyrwich 2007 B [40]	Heart disease	n = 656	Anchor-based, distribution- based and consensus-based	2-15

\*Data from the 1998 National Survey of Functional Health Status (a standard deviation of 23.75883) were used to recalculate norm-based scores to an estimate of the original score. [27]

For anxiety and depression measured by subscales of the Hospital Anxiety and Depression Scale (HADS, score 0-21), the post-treatment differences between the exercise and control group were 1.48 and 1.63 points, respectively. [18] MCID estimates for CFS could not be found, but in patients with chronic obstructive pulmonary disease and cardiovascular disease, MCIDs of 1.5 [41] and 1.7 [42] have been proposed for these scales (range 0.81-5.57, table 3).

**Table 3. Estimates of the minimal clinically important difference (MCID) for Hospital Anxiety and Depression Scale subscales (score 0-21).**

Study	Population	Sample size	Methods	MCID Anxiety subscale (range)	MCID Depression subscale (range)
Puhan 2008 [41]	Chronic obstructive pulmonary disease	n = 88	Anchor-based and distribution-based	1.32-1.57	1.4

Dodd 2011 [43]	Chronic obstructive pulmonary disease	n = 261	Anchor- based	1-1.7	1.3-1.5
Chan 2016 [44]	Survivors of Acute Respiratory Failure	n = 1223	Distribution- based	2.2-2.6	1.9-2.5
Smid 2017 [45]	Chronic obstructive pulmonary disease	n = 419	Anchor- based and distribution- based	1.1-2.0	1.5-1.8
Lemay 2018 [42]	Cardiovascular disease.	n = 591	Anchor- based and distribution- based	0.81-5.21	0.5-5.57

For the outcome of sleep, measured by the Jenkins Sleep Scale, no MCID estimates are available. The authors of the Cochrane review proposed “to view a 20% change in sleep scores as a clinically important difference.” [18] The post-treatment point difference between the exercise and control group was only a 12 percent increase compared to baseline.

These comparisons show that post-treatment differences in exercise trials for CFS are notably small, comparable to (and sometimes lower than) what is considered a minimal meaningful difference.

**2. The risk of response bias was high**

There are several indications that the risk of response bias and expectancy effects was particularly high in exercise trials for CFS.

Exercise therapy was already recommended for CFS by healthcare institutions when the largest trials were being conducted. [46,47] In the largest study, the

investigators published a mid-trial newsletter highlighting that the new NICE guideline recommended GET, among other approaches, “based on the best available evidence” - a strong reinforcement for patients in that treatment arm. [48]

Secondly, patients in the GET group were encouraged to interpret their CFS symptoms as a consequence of deconditioning, stress or anxiety instead of an unknown disease. [49,50] This might have led them to view their disability as more benign compared to patients in the control group who were not primed in this way. According to one therapist manual on GET “participants are encouraged to see symptoms as temporary and reversible, as a result of their current physical weakness, and not as signs of progressive pathology.” [51]

A third indication is that treatment manuals in the randomized trials included strong assertions designed to bolster patients’ expectations of GET. One patient booklet informed trial participants how GET would make them feel: “You will experience a snowballing effect as increasing fitness leads to increasing confidence in your ability. You will have conquered CFS by your own effort and you will be back in control of your body again.” [49] A GET manual for therapists advised regular promotion of the belief that patients can improve: “... it is important that you encourage optimism about the progress that they may make with this approach. You can explain the previous positive research findings of GET and show in the way you discuss goals and use language that you believe they can get better.” [51]

### ***3. Results at short-term follow-up are no longer statistically significant***

Because of close contact with unblinded healthcare providers, assessments made at the end of treatment could be at high risk of response bias. Three large trials of exercise for CFS, which provide most of the data in the Cochrane review, chose to define their primary outcome at follow-up assessment. [8,9,52] One of these trials noted that

“assessment at week 70 is required because short-term assessments of outcome in a chronic health condition such as CFS/ME can be misleading.” [53] Consequently, the findings at follow-up can be seen as an important indication whether reported improvements are robust.

The meta-analyses provided in the Cochrane review show that 52 to 70 weeks after randomization, differences between the exercise and control groups are no longer statistically significant for outcomes of fatigue, physical function, anxiety, depression, as well as pain, quality of life and various measures of healthcare use. [18] The exceptions were specialist medical care contacts and sleep, for which only one and three trials reported data, respectively.

**4. Objective outcomes show little improvement**

Objective outcomes are believed to be generally less prone to bias than PROMs. [54,55] One of the largest studies to date on bias in randomized trials, the BRANDO project, gave the following recommendation: “Our results suggest that, as far as possible, clinical and policy decisions should not be based on trials in which blinding is not feasible and outcome measures are subjectively assessed. Therefore, trials in which blinding is not feasible should focus as far as possible on objectively measured outcomes...” [56]

The eight randomized trials included in the review have data on employment [57,58], disability benefits [58], activity levels [59], and fitness tests [60,61]. A narrative review of these objective outcomes is provided by Vink and Vink-Niese [62] Overall, there were few significant improvements in the exercise group compared to the passive control group. Vink and Vink-Niese conclude that “when the objective outcomes of the trials are considered, it is possible to state that GET is ineffective for CFS.” [62]

### ***5. Mediation analyses do not support assumptions about deconditioning***

The lack of improvement on objective measures of fitness is puzzling given that the explicit rationale for exercise therapy was to treat patients' presumed deconditioning. [10] Patient booklets and instruction manuals used in the randomized trials explain how deconditioning could account for many or all of the symptoms seen in CFS and how exercise therapy would reverse these impairments. [49,51] In the largest trial included in the review, the PACE study, GET "was done on the basis of deconditioning and exercise intolerance theories of chronic fatigue syndrome. These theories assume that the syndrome is perpetuated by reversible physiological changes of deconditioning and avoidance of activity." [8]

These assumptions are not supported by current evidence. The five trials that explored a mediation analysis all found that self-reported improvements in fatigue and physical function were poorly mediated by objective measures of fitness. [60,61,63–65] According to Cheshire et al. (2018) "the exact mechanisms underpinning any improvement following GET are unclear. We know that GET does not work by improving physical fitness, although it may improve exercise tolerance; it may be that it also works as a graded exposure therapy by reducing fear avoidance." [66]

Hypothesized mechanisms for the latter hypotheses have not been adequately articulated in the scientific literature. It is unclear how a reduction in fear-avoidance or exercise intolerance would explain the reported reductions in fatigue or improvements in physical function, given that patients did not significantly increase their activity levels [59], physical fitness [60,61] or employment status [57,58]. Therefore, we conclude that exercise therapy in CFS seems to lack a plausible mechanism for improvement.

**Discussion**

In this study we have argued that treatment effects in exercise trials for CFS are small, comparable to (and sometimes lower than) what is considered a minimal meaningful change, and no longer statistically significant at follow-up assessments. The risk of response bias and expectancy effects was high as patients were actively encouraged to adopt a positive attitude towards GET. The improvements reported on PROMs were not seen on objective measurements believed to be less prone to bias associated with lack of blinding. Mediation analyses do not support the deconditioning hypothesis and indicate that exercise therapy in CFS currently lacks a plausible mechanism for improvement. Based on the arguments outlined above, we conclude that treatment effects seen in exercise trials for CFS are likely the result of bias associated with a lack of blinding.

A recent meta-epidemiological study, the MetaBLIND study, could not find an average difference in estimated treatment effect between trials with and without blinded patients or healthcare providers. [67] One possible explanation for this finding is selection bias: interventions where both blinded and unblinded trials are available could also be those where the impact of blinding is considered less important. We have argued that in exercise trials for CFS the risk of response bias is particularly high because of close contact and collaboration with therapists and active encouragements to maintain an optimistic attitude towards GET.

Some have argued that expectancy effects might have been reduced in exercise trials for CFS because of opposition to GET by patient organizations and online fora. [18] One of the trials mentioned that “recruitment was severely handicapped by the appearance of an Australian article in the lay CFS literature suggesting that exercise therapy for CFS patients was harmful and should be avoided.” [65] Only one exercise trial, the PACE study [8], measured the credibility of interventions before treatment

started. Patients were asked to rate how logical their proposed treatment seemed and how confident they were that it would help them. On both questions, GET (84% and 70%) scored better than cognitive behavioural therapy (71% and 57%) and specialist medical care (49% and 41%), and similar to adaptive pacing therapy (84% and 72%) - the method favored by patients in multiple surveys [13]. One exercise trial for patients with CFS called GETSET was not included in the meta-analyses we used. [21] The trial compared guided graded exercise self-help with specialist medical care (SMC) to SMC alone. Treatments effects were comparable to the mean differences discussed here, namely a 4.2 point difference on the Chalder Fatigue Scale, a 6.3 points difference on the SF-36 physical functioning subscale and 1.2 and 1.1 point differences on the HADS depression and anxiety subscales respectively. Therefore inclusion of this trial would change little to our analysis and conclusion.

We focused our analysis on the comparison of GET to a passive control group as this was the design used by most exercise trials for CFS. It's possible that the small to moderate treatment effects in these trials reflect the lack of an active and credible control group rather than bias associated with a lack of blinding. In the PACE trial, however, GET outperformed Adaptive Pacing Therapy (APT) with mean differences comparable to those discussed here. [8]

We acknowledge a level of uncertainty about our hypothesis but argue that the burden of proof lies, as for all interventions, on those claiming efficacy. Researchers and healthcare agencies who assert that a treatment is effective should provide evidence that reported improvements are not merely the result of bias and methodological weaknesses. We argue that this is currently not the case for exercise therapy in CFS where reported improvements are likely the result of bias associated with a lack of blinding and reliance on subjective outcome measures.

**Conclusions**

To assess the validity of patient-reported improvements in non-blinded trials, we have analyzed their magnitude, consistency over time and congruence with objective measurements. We have reviewed mediation analyses, plausible mechanisms of improvement and the risk of expectancy effects and other biases. We hope this case study furthers critical assessment of patient-reported improvements in areas of medicine where blinding of therapists and trial participants faces practical limitations.

**Notes**

1. A reviewer of this paper drew our attention to the study by Cho et al. [68] which argues that the placebo response is low in the CFS patient population compared to other illnesses. We contest the conclusion of this review on several grounds. Cho and colleagues used a definition of the placebo effect as the response rate in the control arm of randomized clinical trials. The authors then estimated the size of the placebo effect in CFS studies by doing a meta-analysis of the percentage of responders in the control arm of several randomized clinical trials for CFS. This method is problematic because the response rate in the control group can be affected by multiple factors, including the natural progression of the illness. The low response rate that Cho and colleagues interpret as a placebo effect could also reflect a low rate of spontaneous improvement in patients with CFS. Second, Cho et al. also included passive control conditions where patients did not receive an active intervention but were put on a waiting list or received usual care as was the case in the trial by Sharpe et al. 1996 [69] and Powell et al. 2001 [9]. In our view, the placebo response cannot be estimated if patients do not receive a credible or sham intervention. Third, the response rate of the control group is defined differently in each of the included clinical trials, and depends both on the primary outcome, and on the thresholds used to define a significant improvement. Fourth, the



review only includes a limited number of CFS trials (n = 29) all published before August 2002. The review is out of date, and its finding of a lower response rate in the control group of CFS trials compared to other conditions could be due to chance and a small sample size. Recent, large, and well-conducted randomized trials for CFS point to a significant response rate in the control group. The most notable example of this is the phase III trial on rituximab which reported a response rate of 35% in the control group. [70]

### Acknowledgements

The authors would like to thank Tom Kindlon, Simon McGrath, and Andrew Kewley for their thoughtful comments to earlier drafts of this analysis.

### Competing interests

MT is the assistant chairperson of the Belgian patient organization 12ME. Members of the ME/CFS patient community donated to a crowdfunding campaign in support of DMT's academic position. CS reports no conflict of interest.

### Funding

None

### Authors' contributions

MT conceived the idea for the analysis and wrote the initial version. DMT and CS refined the arguments and structure of the manuscript and made substantial additions to the first and subsequent drafts. All authors read and approved the final manuscript.

### References

- [1] Nacul LC, Lacerda EM, Pheby D, et al. Prevalence of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) in three regions of England: a repeated cross-sectional study in primary care. *BMC Med.* 2011;9:91.
- [2] Jason LA, Richman JA, Rademaker AW, et al. A community-based study of chronic fatigue syndrome. *Arch Intern Med.* 1999;159:2129–2137.

- [3] Fukuda K, Straus SE, Hickie I, et al. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med.* 1994;121:953–959.
- [4] Institute of Medicine. Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness [Internet]. Washington (DC): National Academies Press (US); 2015. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK274235/>.
- [5] Edwards JCW, McGrath S, Baldwin A, et al. The biological challenge of myalgic encephalomyelitis/chronic fatigue syndrome: a solvable problem. *Fatigue Biomed Health Behav.* 2016;4:63–69.
- [6] Smith MEB, Nelson HD, Haney E, et al. Diagnosis and Treatment of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Evid Report Technology Assess.* 2014;1–433.
- [7] Carruthers BM, Sande MI van de, Meirleir KLD, et al. Myalgic encephalomyelitis: International Consensus Criteria. *J Intern Med.* 2011;270:327–338.
- [8] White PD, Goldsmith KA, Johnson AL, et al. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. *Lancet Lond Engl.* 2011;377:823–836.
- [9] Powell P, Bentall RP, Nye FJ, et al. Randomised controlled trial of patient education to encourage graded exercise in chronic fatigue syndrome. *BMJ.* 2001;322:387–390.
- [10] Clark LV, White PD. The role of deconditioning and therapeutic exercise in chronic fatigue syndrome (CFS). *J Ment Health.* 2005;14:237–252.
- [11] ME Association. Petition: The NICE guideline for CFS/ME is unfit for purpose and needs a complete revision. [Internet]. Change.org. [cited 2019 Nov 26]. Available from: <https://www.change.org/p/petition-the-nice-guideline-for-cfs-me-is-unfit-for-purpose-and-needs-a-complete-revision>.
- [12] Kindlon T. Reporting of Harms Associated with Graded Exercise Therapy and Cognitive Behavioural Therapy in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Bull IACFSME.* 2011;19:59–111.
- [13] Geraghty K, Hann M, Kurtev S. Myalgic encephalomyelitis/chronic fatigue syndrome patients' reports of symptom changes following cognitive behavioural therapy, graded exercise therapy and pacing treatments: Analysis of a primary survey compared with secondary surveys. *J Health Psychol.* 2019;24:1318–1333.
- [14] Edwards J. PACE team response shows a disregard for the principles of science. *J Health Psychol.* 2017;22:1155–1158.

- [15] Wilshire CE, Kindlon T, Courtney R, et al. Rethinking the treatment of chronic fatigue syndrome-a reanalysis and evaluation of findings from a recent major trial of graded exercise and CBT. *BMC Psychol*. 2018;6:6.
- [16] Hróbjartsson A, Boutron I. Blinding in randomized clinical trials: imposed impartiality. *Clin Pharmacol Ther*. 2011;90:732–736.
- [17] Kamper SJ. Blinding: Linking Evidence to Practice. *J Orthop Sports Phys Ther*. 2018;48:825–826.
- [18] Larun L, Brurberg KG, Odgaard-Jensen J, et al. Exercise therapy for chronic fatigue syndrome. Cochrane Common Mental Disorders Group, editor. *Cochrane Database Syst Rev* [Internet]. 2019 [cited 2019 Nov 26]; Available from: <http://doi.wiley.com/10.1002/14651858.CD003200.pub8>.
- [19] Smid DE, Franssen FME, Houben-Wilke S, et al. Responsiveness and MCID Estimates for CAT, CCQ, and HADS in Patients With COPD Undergoing Pulmonary Rehabilitation: A Prospective Analysis. *J Am Med Dir Assoc*. 2017;18:53–58.
- [20] Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, et al. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol*. 2014;43:1272–1283.
- [21] Clark LV, Pesola F, Thomas JM, et al. Guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome (GETSET): a pragmatic randomised controlled trial. *The Lancet*. 2017;390:363–373.
- [22] Ridsdale L, Godfrey E, Seed P. Chronic Fatigue in general practice: authors reply. *Br J Gen Pract*. 2001;51:317–318.
- [23] Goligher EC, Pouchot J, Brant R, et al. Minimal clinically important difference for 7 measures of fatigue in patients with systemic lupus erythematosus. *J Rheumatol*. 2008;35:635–642.
- [24] Pouchot J, Kherani RB, Brant R, et al. Determination of the minimal clinically important difference for seven fatigue measures in rheumatoid arthritis. *J Clin Epidemiol*. 2008;61:705–713.
- [25] Pettersson S, Lundberg IE, Liang MH, et al. Determination of the minimal clinically important difference for seven measures of fatigue in Swedish patients with systemic lupus erythematosus. *Scand J Rheumatol*. 2015;44:206–210.
- [26] Nøstdahl T, Bernklev T, Fredheim OM, et al. Defining the cut-off point of clinically significant postoperative fatigue in three common fatigue scales. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2019;28:991–1003.
- [27] Ware JE, Kosinski M, Dewey JE. How to score version 2 of the SF-36 health survey: standards & acute forms. Lincoln, R.I.: QualityMetric; 2001.

- [28] Clark LV, McCrone P, Ridge D, et al. Graded Exercise Therapy Guided Self-Help Trial for Patients with Chronic Fatigue Syndrome (GETSET): Protocol for a Randomized Controlled Trial and Interview Study. *JMIR Res Protoc*. 2016;5:e70.
- [29] Brigden A, Parslow RM, Gaunt D, et al. Defining the minimally clinically important difference of the SF-36 physical function subscale for paediatric CFS/ME: triangulation using three different methods. *Health Qual Life Outcomes*. 2018;16:202.
- [30] Kosinski M, Zhao SZ, Dedhiya S, et al. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum*. 2000;43:1478–1487.
- [31] Ward MM, Guthrie LC, Alba MI. Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis Care Res*. 2014;66:1783–1789.
- [32] Hays RD, Spritzer KL, Fries JF, et al. Responsiveness and minimally important difference for the patient-reported outcomes measurement information system (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. *Ann Rheum Dis*. 2015;74:104–107.
- [33] de Groot V, Beckerman H, Uitdehaag BMJ, et al. The usefulness of evaluative outcome measures in patients with multiple sclerosis. *Brain J Neurol*. 2006;129:2648–2659.
- [34] Robinson D, Zhao N, Gathany T, et al. Health perceptions and clinical characteristics of relapsing-remitting multiple sclerosis patients: baseline data from an international clinical trial. *Curr Med Res Opin*. 2009;25:1121–1130.
- [35] Devilliers H, Amoura Z, Besancenot J-F, et al. Responsiveness of the 36-item Short Form Health Survey and the Lupus Quality of Life questionnaire in SLE. *Rheumatol Oxf Engl*. 2015;54:940–949.
- [36] McElhone K, Abbott J, Sutton C, et al. Sensitivity to Change and Minimal Important Differences of the LupusQoL in Patients With Systemic Lupus Erythematosus. *Arthritis Care Res*. 2016;68:1505–1513.
- [37] Coteur G, Feagan B, Keininger DL, et al. Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. *Aliment Pharmacol Ther*. 2009;29:1032–1041.
- [38] Erez G, Selman L, Murtagh FEM. Measuring health-related quality of life in patients with conservatively managed stage 5 chronic kidney disease: limitations of the Medical Outcomes Study Short Form 36: SF-36. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2016;25:2799–2809.
- [39] Wyrwich KW, Metz SM, Kroenke K, et al. Measuring patient and clinician perspectives to evaluate change in health-related quality of life among patients with chronic obstructive pulmonary disease. *J Gen Intern Med*. 2007;22:161–170.

- [40] Wyrwich KW, Metz SM, Kroenke K, et al. Triangulating patient and clinician perspectives on clinically important differences in health-related quality of life among patients with heart disease. *Health Serv Res*. 2007;42:2257–2274; discussion 2294–2323.
- [41] Puhan MA, Frey M, Büchi S, et al. The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. *Health Qual Life Outcomes*. 2008;6:46.
- [42] Lemay KR, Tulloch HE, Pipe AL, et al. Establishing the Minimal Clinically Important Difference for the Hospital Anxiety and Depression Scale in Patients With Cardiovascular Disease. *J Cardiopulm Rehabil Prev*. 2019;39:E6–E11.
- [43] Dodd JW, Hogg L, Nolan J, et al. The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicentre, prospective study. *Thorax*. 2011;66:425–429.
- [44] Chan KS, Aronson Friedman L, Bienvenu OJ, et al. Distribution-based estimates of minimal important difference for hospital anxiety and depression scale and impact of event scale-revised in survivors of acute respiratory failure. *Gen Hosp Psychiatry*. 2016;42:32–35.
- [45] Smid DE, Franssen FME, Houben-Wilke S, et al. Responsiveness and MCID Estimates for CAT, CCQ, and HADS in Patients With COPD Undergoing Pulmonary Rehabilitation: A Prospective Analysis. *J Am Med Dir Assoc*. 2017;18:53–58.
- [46] CFS/ME Working Group. Report to the Chief Medical Officer of an independent working group. [Internet]. Department of Health; 2002 [cited 2019 Dec 13]. Available from: <https://www.meassociation.org.uk/wp-content/uploads/CMO-Report-2002.pdf>.
- [47] Working Group convened under the auspices of, the Royal Australasian College of Physicians. Chronic fatigue syndrome. Clinical practice guidelines. [Internet]. Health Policy Unit, Royal Australasian College of Physicians; 2002 [cited 2019 Nov 26]. Available from: [https://www.mja.com.au/system/files/issues/cfs2\\_2.pdf](https://www.mja.com.au/system/files/issues/cfs2_2.pdf).
- [48] Tuller D. TRIAL BY ERROR: The Troubling Case of the PACE Chronic Fatigue Syndrome Study. (second installment). [Internet]. Virol. Blog. 2015 [cited 2019 Nov 26]. Available from: <http://www.virology.ws/2015/10/22/trial-by-error-ii/>.
- [49] Powell P. FINE trial patient booklet version 9. [Internet]. Royal Liverpool University Hospital; 2005 [cited 2019 Dec 28]. Available from: <https://web.archive.org/web/20140811161130/http://www.fine-trial.net/downloads/Patient%20PR%20Manual%20ver9%20Apr05.pdf>.
- [50] Bavinton J, Dyer N, White PD. PACE. Graded exercise therapy, information for participants. Final trial version: (MREC version 1). [Internet]. PACE Trial Management Group; 2004 [cited 2019 Dec 28]. Available from: <https://media.org/images/d/d4/PACE-get-participant-manual.pdf>.

- [51] Bavinton J, Darbishire L, White PD. PACE manual for therapists. Graded Exercise Therapy (GET) for CFS/ME. Final trial version: version 7 (MREC Version 2). [Internet]. PACE Trial Management Group; 2004 [cited 2019 Nov 26]. Available from: <https://me-pedia.org/images/8/89/PACE-get-therapist-manual.pdf>.
- [52] Wearden AJ, Dowrick C, Chew-Graham C, et al. Nurse led, home based self help treatment for patients in primary care with chronic fatigue syndrome: randomised controlled trial. *BMJ*. 2010;340:c1777.
- [53] Wearden AJ, Riste L, Dowrick C, et al. Fatigue Intervention by Nurses Evaluation--the FINE Trial. A randomised controlled trial of nurse led self-help treatment for patients in primary care with chronic fatigue syndrome: study protocol. [ISRCTN74156610]. *BMC Med*. 2006;4:9.
- [54] Forbes D. Blinding: an essential component in decreasing risk of bias in experimental designs. *Evid Based Nurs*. 2013;16:70–71.
- [55] Kahan BC, Cro S, Doré CJ, et al. Reducing bias in open-label trials where blinded outcome assessment is not feasible: strategies from two randomised trials. *Trials*. 2014;15:456.
- [56] Savović J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess Winch Engl*. 2012;16:1–82.
- [57] Jason LA, Torres-Harding S, Friedberg F, et al. Non-pharmacologic Interventions for CFS: A Randomized Trial. *J Clin Psychol Med Settings*. 2007;14:275–296.
- [58] McCrone P, Sharpe M, Chalder T, et al. Adaptive Pacing, Cognitive Behaviour Therapy, Graded Exercise, and Specialist Medical Care for Chronic Fatigue Syndrome: A Cost-Effectiveness Analysis. *PLOS ONE*. 2012;7:e40808.
- [59] Wallman KE, Morton AR, Goodman C, et al. Randomised controlled trial of graded exercise in chronic fatigue syndrome. *Med J Aust*. 2004;180:444–448.
- [60] Wearden AJ, Emsley R. Mediators of the effects on fatigue of pragmatic rehabilitation for chronic fatigue syndrome. *J Consult Clin Psychol*. 2013;81:831–838.
- [61] Chalder T, Goldsmith KA, White PD, et al. Rehabilitative therapies for chronic fatigue syndrome: a secondary mediation analysis of the PACE trial. *Lancet Psychiatry*. 2015;2:141–152.
- [62] Vink M, Vink-Niese A. Graded exercise therapy for myalgic encephalomyelitis/chronic fatigue syndrome is not effective and unsafe. Re-analysis of a Cochrane review. *Health Psychol Open*. 2018;5:2055102918805187.
- [63] Fulcher KY, White PD. Randomised controlled trial of graded exercise in patients with the chronic fatigue syndrome. *BMJ*. 1997;314:1647–1652.



- [64] Wearden AJ, Morriss RK, Mullis R, et al. Randomised, double-blind, placebo-controlled treatment trial of fluoxetine and graded exercise for chronic fatigue syndrome. *Br J Psychiatry J Ment Sci.* 1998;172:485–490.
- [65] Moss-Morris R, Sharon C, Tobin R, et al. A randomized controlled graded exercise trial for chronic fatigue syndrome: outcomes and mechanisms of change. *J Health Psychol.* 2005;10:245–259.
- [66] Cheshire A, Ridge D, Clark L, et al. Guided graded Exercise Self-help for chronic fatigue syndrome: patient experiences and perceptions. *Disabil Rehabil.* 2018;1–10.
- [67] Moustgaard H, Clayton GL, Jones HE, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ.* 2020;368:l6802.
- [68] Cho HJ, Hotopf M, Wessely S. The placebo response in the treatment of chronic fatigue syndrome: a systematic review and meta-analysis. *Psychosom Med.* 2005;67:301–313.
- [69] Sharpe M, Hawton K, Simkin S, et al. Cognitive behaviour therapy for the chronic fatigue syndrome: a randomized controlled trial. *BMJ.* 1996;312:22–26.
- [70] Fluge Ø, Rekeland IG, Lien K, et al. B-Lymphocyte Depletion in Patients With Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial. *Ann Intern Med.* 2019;170:585–593.

**Supplementary file: search strategy**

**Chalder Fatigue Scale**

***Search strategy***

(minimal clinically important difference OR minimal clinically important differences OR “subjectively significant difference” OR “subjectively significant differences” OR “minimal important difference” OR “minimal important differences” OR “clinically important difference” OR “clinically important differences” OR “clinically significant change” OR “clinically significant changes” OR “clinically significant” OR “minimally important change” OR “minimally important changes” OR “minimally detectable difference” OR “minimally detectable differences” OR “minimum important difference” OR “minimum important differences” OR “minimal clinically important change” OR “minimal clinically important changes” OR “minimally important difference” OR “minimally important differences” OR “clinically important change” OR “clinically important changes” OR “minimal clinical important difference” OR “minimal clinical important differences” OR “clinical important differences” OR “clinical important difference” OR “clinical important changes” OR “clinical important change” OR “minimally important clinical difference” OR “minimally important clinical differences” OR “clinically important improvement” OR “clinically important improvements” OR “clinical important improvement” OR “clinical important improvements” OR MID OR MIDs OR MCID OR MCIDs OR CID OR MCIC) AND (Chalder Fatigue Scale OR Chalder Fatigue Questionnaire OR “Fatigue Scale” OR “Fatigue Scales” OR “Fatigue Questionnaire” OR “Fatigue Questionnaires” OR CFS)



### ***Selection of studies***

On 12 November 2019, this search resulted in 163 highlighted publications on PubMed.

Selection after abstract scanning resulted in 15 studies. Selection after full-text scanning resulted in 4 studies. In addition, one study (Ridsdale et al. 2001) was identified from hand searching references, resulting in 5 selected studies in total.

### ***List of studies***

1. Goligher EC, Pouchot J, Brant R, Kherani RB, Aviña-Zubieta JA, Lacaille D, et al. Minimal clinically important difference for 7 measures of fatigue in patients with systemic lupus erythematosus. *J Rheumatol*. 2008;35:635–42.
2. Nøstdahl T, Bernklev T, Fredheim OM, Paddison JS, Raeder J. Defining the cut-off point of clinically significant postoperative fatigue in three common fatigue scales. *Qual Life Res*. 2019;28:991–1003.
3. Pettersson S, Lundberg IE, Liang MH, Pouchot J, Henriksson EW. Determination of the minimal clinically important difference for seven measures of fatigue in Swedish patients with systemic lupus erythematosus. *Scand J Rheumatol*. 2015;44:206–10.
4. Pouchot J, Kherani RB, Brant R, Lacaille D, Lehman AJ, Ensworth S, et al. Determination of the minimal clinically important difference for seven fatigue measures in rheumatoid arthritis. *J Clin Epidemiol*. 2008;61:705–13.
5. Ridsdale L, Godfrey E, Seed P. Chronic Fatigue in general practice: authors reply. *British Journal of General Practice*. 2001;51:317–8.

**Hospital Anxiety and Depression Scale**

***Search strategy***

(minimal clinically important difference OR minimal clinically important differences OR “subjectively significant difference” OR “subjectively significant differences” OR “minimal important difference” OR “minimal important differences” OR “clinically important difference” OR “clinically important differences” OR “clinically significant change” OR “clinically significant changes” OR “clinically significant” OR “minimally important change” OR “minimally important changes” OR “minimally detectable difference” OR “minimally detectable differences” OR “minimum important difference” OR “minimum important differences” OR “minimal clinically important change” OR “minimal clinically important changes” OR “minimally important difference” OR “minimally important differences” OR “clinically important change” OR “clinically important changes” OR “minimal clinical important difference” OR “minimal clinical important differences” OR “clinical important differences” OR “clinical important difference” OR “clinical important changes” OR “clinical important change” OR “minimally important clinical difference” OR “minimally important clinical differences” OR “clinically important improvement” OR “clinically important improvements” OR “clinical important improvement” OR “clinical important improvements” OR “Responsiveness” OR MID OR MIDs OR MCID OR MCIDs OR CID OR MCIC) AND (“Hospital Anxiety and Depression” OR HADS OR HAD)

***Selection of studies***

On 12 November 2019, this search resulted in 406 highlighted publications on PubMed. Selection after abstract scanning resulted in 23 studies. Selection after full-text scanning resulted in 5 studies.

### ***List of studies***

1. Chan KS, Aronson Friedman L, Bienvenu OJ, Dinglas VD, Cuthbertson BH, Porter R, et al. Distribution-based estimates of minimal important difference for hospital anxiety and depression scale and impact of event scale-revised in survivors of acute respiratory failure. *Gen Hosp Psychiatry*. 2016;42:32–5.
2. Dodd JW, Hogg L, Nolan J, Jefford H, Grant A, Lord VM, et al. The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicentre, prospective study. *Thorax*. 2011;66:425–9.
3. Lemay KR, Tulloch HE, Pipe AL, Reed JL. Establishing the Minimal Clinically Important Difference for the Hospital Anxiety and Depression Scale in Patients With Cardiovascular Disease. *J Cardiopulm Rehabil Prev*. 2019;39:E6–11.
4. Puhan MA, Frey M, Büchi S, Schünemann HJ. The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. *Health Qual Life Outcomes*. 2008;6:46.
5. Smid DE, Franssen FME, Houben-Wilke S, Vanfleteren LEGW, Janssen DJA, Wouters EFM, et al. Responsiveness and MCID Estimates for CAT, CCQ, and HADS in Patients With COPD Undergoing Pulmonary Rehabilitation: A Prospective Analysis. *J Am Med Dir Assoc*. 2017;18:53–8.

### **SF-36 physical function subscale**

### ***Search strategy***

(minimal clinically important difference OR minimal clinically important differences OR “subjectively significant difference” OR “subjectively significant differences” OR “minimal important difference” OR “minimal important differences” OR “clinically important difference”

OR “clinically important differences” OR “clinically significant change” OR “clinically significant changes” OR “minimally important change” OR “minimally important changes” OR “minimally detectable difference” OR “minimally detectable differences” OR “minimum important difference” OR “minimum important differences” OR “minimal clinically important change” OR “minimal clinically important changes” OR “minimally important difference” OR “minimally important differences” OR “clinically important change” OR “clinically important changes” OR “minimal clinical important difference” OR “minimal clinical important differences” OR “clinical important differences” OR “clinical important difference” OR “clinical important changes” OR “clinical important change” OR “minimally important clinical difference” OR “minimally important clinical differences” OR “clinically important improvement” OR “clinically important improvements” OR “clinical important improvement” OR “clinical important improvements” OR MID OR MIDs OR MCID OR MCIDs OR CID OR MCIC) AND (SF-36 OR short-Form 36 OR 36-Item short form survey OR short form health survey OR short form health questionnaire OR medical Outcomes Study-Short Form-36 OR MOS-36 OR SF-36v2 OR “physical functioning scale”)

***Selection of studies***

On 11 November 2019, this search resulted in 1268 highlighted publications on PubMed. Selection after abstract scanning resulted in 141 studies. Selection after full-text scanning resulted in 43 studies. In addition, one study (Devilliers 2015) was identified from hand searching references, resulting in 44 selected studies in total.

***List of studies***

1. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using

- 1  
2  
3 WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the  
4 lower extremities. *Arthritis Rheum.* 2001;45:384–91.
- 5  
6  
7  
8 2. Angst F, Benz T, Lehmann S, Aeschlimann A, Angst J. Multidimensional minimal clinically  
9 important differences in knee osteoarthritis after comprehensive rehabilitation: a prospective  
10 evaluation from the Bad Zurzach Osteoarthritis Study. *RMD Open.* 2018;4:e000685.
- 11  
12  
13  
14 3. Antonescu I, Scott S, Tran TT, Mayo NE, Feldman LS. Measuring postoperative recovery:  
15 what are clinically meaningful differences? *Surgery.* 2014;156:319–27.
- 16  
17  
18  
19 4. Brigden A, Parslow RM, Gaunt D, Collin SM, Jones A, Crawley E. Defining the minimally  
20 clinically important difference of the SF-36 physical function subscale for paediatric CFS/ME:  
21 triangulation using three different methods. *Health Qual Life Outcomes.* 2018;16:202.
- 22  
23  
24  
25 5. Cabanas-Grandío P, García-Seara J, Gude F, Martínez-Sande JL, Fernández-López XA,  
26 González-Juanatey JR. Assessment of long-term quality of life after cavotricuspid isthmus  
27 ablation for typical atrial flutter. *Health Qual Life Outcomes.* 2014;12:47.
- 28  
29  
30  
31  
32 6. Campos WK, Linhares MN, Sarda J, Santos ARS, Licinio J, Quevedo J, et al. Determinants for  
33 Meaningful Clinical Improvement of Pain and Health-Related Quality of Life After Spinal Cord  
34 Stimulation for Chronic Intractable Pain. *Neuromodulation: Technology at the Neural Interface.*  
35 2019;22:280–9.
- 36  
37  
38  
39 7. Coe MP, Sutherland JM, Penner MJ, Younger A, Wing KJ. Minimal clinically important  
40 difference and the effect of clinical variables on the ankle osteoarthritis scale in surgically treated  
41 end-stage ankle arthritis. *J Bone Joint Surg Am.* 2015;97:818–23.
- 42  
43  
44  
45 8. Coteur G, Feagan B, Keininger DL, Kosinski M. Evaluation of the meaningfulness of health-  
46 related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients  
47 with active Crohn's disease. *Aliment Pharmacol Ther.* 2009;29:1032–41.
- 48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

9. Davidson M, Keating JL, Eyres S. A low back-specific version of the SF-36 Physical Functioning scale. *Spine*. 2004;29:586–94.
10. Dawson J, Doll H, Coffey J, Jenkinson C, Oxford and Birmingham Foot and Ankle Clinical Research Group. Responsiveness and minimally important change for the Manchester-Oxford foot questionnaire (MOXFQ) compared with AOFAS and SF-36 assessments following surgery for hallux valgus. *Osteoarthr Cartil*. 2007;15:918–31.
11. Dawson J, Boller I, Doll H, Lavis G, Sharp R, Cooke P, et al. Minimally important change was estimated for the Manchester-Oxford Foot Questionnaire after foot/ankle surgery. *J Clin Epidemiol*. 2014;67:697–705.
12. de Groot V, Beckerman H, Uitdehaag BMJ, de Vet HCW, Lankhorst GJ, Polman CH, et al. The usefulness of evaluative outcome measures in patients with multiple sclerosis. *Brain*. 2006;129 Pt 10:2648–59.
13. Devilliers H, Amoura Z, Besancenot J-F, Bonnotte B, Pasquali J-L, Wahl D, et al. Responsiveness of the 36-item Short Form Health Survey and the Lupus Quality of Life questionnaire in SLE. *Rheumatology (Oxford)*. 2015;54:940–9.
14. Erez G, Selman L, Murtagh FEM. Measuring health-related quality of life in patients with conservatively managed stage 5 chronic kidney disease: limitations of the Medical Outcomes Study Short Form 36: SF-36. *Qual Life Res*. 2016;25:2799–809.
15. Escobar A, Quintana JM, Bilbao A, Aróstegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. *Osteoarthr Cartil*. 2007;15:273–80.
16. Gifford JM, Husain N, Dinglas VD, Colantuoni E, Needham DM. Baseline quality of life before intensive care: a comparison of patient versus proxy responses. *Crit Care Med*. 2010;38:855–60.

17. Gilbert C, Brown MCJ, Cappelleri JC, Carlsson M, McKenna SP. Estimating a minimally important difference in pulmonary arterial hypertension following treatment with sildenafil. *Chest*. 2009;135:137–42.
18. Hays RD, Spritzer KL, Fries JF, Krishnan E. Responsiveness and minimally important difference for the patient-reported outcomes measurement information system (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. *Ann Rheum Dis*. 2015;74:104–7.
19. Jayadevappa R, Malkowicz SB, Wittink M, Wein AJ, Chhatre S. Comparison of distribution- and anchor-based approaches to infer changes in health-related quality of life of prostate cancer survivors. *Health Serv Res*. 2012;47:1902–25.
20. Jones G, Jenkinson C, Kennedy S. Evaluating the responsiveness of the Endometriosis Health Profile Questionnaire: the EHP-30. *Qual Life Res*. 2004;13:705–13.
21. Keurentjes JC, Fiocco M, Nelissen RG. Willingness to undergo surgery again validated clinically important differences in health-related quality of life after total hip replacement or total knee replacement surgery. *J Clin Epidemiol*. 2014;67:114–20.
22. Khatri D, Jaiswal A, Das KK, Pandey S, Bhaisora K, Kumar R. Health-related quality of life after surgery in supratentorial gliomas. *Neurol India*. 2019;67:467–75.
23. Knox SA, King MT. Validation and calibration of the SF-36 health transition question against an external criterion of clinical change in health status. *Qual Life Res*. 2009;18:637–45.
24. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum*. 2000;43:1478–87.
25. Lee AC, Driban JB, Price LL, Harvey WF, Rodday AM, Wang C. Responsiveness and Minimally Important Differences for 4 Patient-Reported Outcomes Measurement Information

System Short Forms: Physical Function, Pain Interference, Depression, and Anxiety in Knee Osteoarthritis. *J Pain*. 2017;18:1096–110.

26. Leung Y-Y, Zhu TY, Tam LS, Kun EW-L, Li EK-M. Minimal important difference and responsiveness to change of the SF-36 in patients with psoriatic arthritis receiving tumor necrosis factor- $\alpha$  blockers. *J Rheumatol*. 2011;38:2077–9.

27. McElhone K, Abbott J, Sutton C, Mullen M, Lanyon P, Rahman A, et al. Sensitivity to Change and Minimal Important Differences of the LupusQoL in Patients With Systemic Lupus Erythematosus. *Arthritis Care Res (Hoboken)*. 2016;68:1505–13.

28. Pérez-Fentes DA, Gude F, Blanco B, Freire CG. Percutaneous nephrolithotomy: short- and long-term effects on health-related quality of life. *J Endourol*. 2015;29:13–7.

29. Quintana JM, Escobar A, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. *Osteoarthr Cartil*. 2005;13:1076–83.

30. Robinson D, Zhao N, Gathany T, Kim L-L, Cella D, Revicki D. Health perceptions and clinical characteristics of relapsing-remitting multiple sclerosis patients: baseline data from an international clinical trial. *Curr Med Res Opin*. 2009;25:1121–30.

31. Shi H-Y, Chang J-K, Wong C-Y, Wang J-W, Tu Y-K, Chiu H-C, et al. Responsiveness and minimal important differences after revision total hip arthroplasty. *BMC Musculoskelet Disord*. 2010;11:261.

32. Spratt KF. Patient-level minimal clinically important difference based on clinical judgment and minimally detectable measurement difference: a rationale for the SF-36 physical function scale in the SPORT intervertebral disc herniation cohort. *Spine*. 2009;34:1722–31.

33. Swigris JJ, Brown KK, Behr J, du Bois RM, King TE, Raghu G, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respir Med*. 2010;104:296–304.



34. Ward MM, Guthrie LC, Alba MI. Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis Care Res (Hoboken)*. 2014;66:1783–9.
35. Witt S, Krauss E, Barbero MAN, Müller V, Bonniaud P, Vancheri C, et al. Psychometric properties and minimal important differences of SF-36 in Idiopathic Pulmonary Fibrosis. *Respir Res*. 2019;20:47.
36. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care*. 1999;37:469–78.
37. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999;52:861–73.
38. Wyrwich KW, Fihn SD, Tierney WM, Kroenke K, Babu AN, Wolinsky FD. Clinically important changes in health-related quality of life for patients with chronic obstructive pulmonary disease: an expert consensus panel report. *J Gen Intern Med*. 2003;18:196–202.
39. Wyrwich KW, Metz SM, Kroenke K, Tierney WM, Babu AN, Wolinsky FD. Interpreting quality-of-life data: methods for community consensus in asthma. *Ann Allergy Asthma Immunol*. 2006;96:826–33.
40. Wyrwich KW, Metz SM, Kroenke K, Tierney WM, Babu AN, Wolinsky FD. Measuring patient and clinician perspectives to evaluate change in health-related quality of life among patients with chronic obstructive pulmonary disease. *J Gen Intern Med*. 2007;22:161–70.
41. Wyrwich KW, Metz SM, Kroenke K, Tierney WM, Babu AN, Wolinsky FD. Triangulating patient and clinician perspectives on clinically important differences in health-related quality of

life among patients with heart disease. Health Serv Res. 2007;42 6 Pt 1:2257–74; discussion 2294-2323.

42. Wyrwich KW, Nelson HS, Tierney WM, Babu AN, Kroenke K, Wolinsky FD. Clinically important differences in health-related quality of life for patients with asthma: an expert consensus panel report. Ann Allergy Asthma Immunol. 2003;91:148–53.

43. Wyrwich KW, Spertus JA, Kroenke K, Tierney WM, Babu AN, Wolinsky FD, et al. Clinically important differences in health status for patients with heart disease: an expert consensus panel report. Am Heart J. 2004;147:615–22.

44. Wyrwich KW, Tierney WM, Babu AN, Kroenke K, Wolinsky FD. A comparison of clinically important differences in health-related quality of life for patients with chronic lung disease, asthma, or heart disease. Health Serv Res. 2005;40:577–91.

**Second selection**

A full overview of the 44 MCID estimates and the methods used in each of these studies, some of which have a little relevance to CFS, was considered outside the scope of this paper. Therefore, the authors selected a brief overview of MCID estimates in chronic conditions that are comparable to CFS. These selected estimates were similar to the 44 MCID estimates as a whole.

**Jenkins Sleep Scale**

**Search strategy**

(minimal clinically important difference OR minimal clinically important differences OR “subjectively significant difference” OR “subjectively significant differences” OR “minimal important difference” OR “minimal important differences” OR “clinically important difference”

OR “clinically important differences” OR “clinically significant change” OR “clinically significant changes” OR “clinically significant” OR “minimally important change” OR “minimally important changes” OR “minimally detectable difference” OR “minimally detectable differences” OR “minimum important difference” OR “minimum important differences” OR “minimal clinically important change” OR “minimal clinically important changes” OR “minimally important difference” OR “minimally important differences” OR “clinically important change” OR “clinically important changes” OR “minimal clinical important difference” OR “minimal clinical important differences” OR “clinical important differences” OR “clinical important difference” OR “clinical important changes” OR “clinical important change” OR “minimally important clinical difference” OR “minimally important clinical differences” OR “clinically important improvement” OR “clinically important improvements” OR “clinical important improvement” OR “clinical important improvements” OR “Responsiveness” OR MID OR MIDs OR MCID OR MCIDs OR CID OR MCIC) AND (Jenkins Sleep Questionnaire OR Jenkins Sleep Scale OR JSQ OR JSS OR JSEQ)

### ***Selection of studies***

On 13 November 2019, this resulted in 67 highlighted publications on PubMed, none of which were found to be relevant to the scope of our article.