REVIEW

A systematic review of outcomes of chronic disease self-management interventions

Sandra Nolte · Richard H. Osborne

Abstract

Purpose To carry out a systematic review of program outcomes used in the evaluation of group-based self-management interventions aimed at people with arthritis and other chronic conditions.

Methods The systematic search was performed across databases MEDLINE, EMBASE, CINAHL, and PsycINFO. Both between-group and within-group effect sizes (ES) were calculated. Results were interpreted as small (ES ∼ 0.2), medium (ES ∼ 0.5), or large (ES ∼ 0.8) effects.

Results The majority of 18 included trials investigated the effectiveness of arthritis-specific interventions. Across most outcomes, small effects on course participants were shown. While effects on knowledge were large (between-group ES = 0.78), effects on clinical outcomes such as pain (ES = 0.10) were negligible to small.

Conclusions This paper is consistent with other reviews in this area, suggesting that people with arthritis receive only marginal benefits from participating in chronic disease self-management interventions. When looking at the types of outcomes that trials are based on, however, alternative explanations for these results seem probable. As evaluations heavily rely on patient self-report, current approaches to program evaluation may not be sufficient to assess the intended impact of self-management education. An in-depth investigation of the types of outcomes assessed is provided in a separate paper.

Introduction

The prevalence of chronic disease has risen dramatically over past decades, making it the major cause of deaths worldwide [1]. In 2005, the World Health Organization estimated that chronic illnesses accounted for 49 % of the total global burden of disease, and when basing the analysis on people aged 30 years and above, this proportion was estimated to be as high as 72 % [2, 3]. To respond to this burden and reduce both the societal and the economic impact of chronic illnesses, effective chronic disease management is essential. While disease management programs are systematic approaches to navigate patients through the healthcare system and improve quality of care [4], successful management of a chronic illness is highly dependent on the individual patients who have to take extensive responsibility themselves [5]; that is, as affected persons spend the majority of the time outside the healthcare system, they have to learn to manage their chronic disease on their own in their own time. Therefore, active self-management and interventions supporting patients in the acquisition of skills and techniques to learn to live with their disease are key components in managing a chronic condition.

While the logic of providing self-management education for people with chronic disease is clear, it is surprising that current evidence of effectiveness of these interventions is rather mixed. A summary of meta-analytic and systematic

S. Nolte (✉)
Medical Clinic for Psychosomatics, Charité, University Medicine Berlin, Luisenstraße 13a, 10117 Berlin, Germany
e-mail: sandra.nolte@charite.de

S. Nolte · R. H. Osborne
Public Health Innovation, Population Health Strategic Research Centre, School of Health and Social Development, Faculty of Health, Deakin University, Burwood, VIC, Australia

reviews on 'self-management', 'psycho-educational', or 'psychological' programs suggests that there is reasonable evidence that these interventions are beneficial for a wide range of people and courses are regarded as an important adjunct to standard medical care [6–10]. For example, individuals with diabetes have been reported to receive small effects in outcomes such as fasting blood glucose levels and medium effects in both glycated hemoglobin (HbA_{1c}) and psychological variables [7, 9, 11]. Reviews on programs for hypertension have shown positive effects on systolic blood pressure [7, 11]. In contrast, in studies with people with arthritis, fewer benefits have been reported. Despite this group receiving much attention in the literature, most studies report negligible to small effects in outcomes such as disability, function, impairment, and pain [6–8, 11, 12]. Similarly, a narrative review shows that only few trials report positive effects for disability, pain, painful and swollen joints, and symptoms [5].

In Australia, several programs for chronic disease self-management education are offered in a variety of settings. Despite the large number of different interventions, ranging from face-to-face consultations to multimedia campaigns [13], group-based courses are the most common form of self-management program delivery [14]. Around 50 organizations, including most Australian Arthritis Foundations, are currently licensed to run self-management courses following the curricula of the Stanford Patient Education Research Center [15]. These courses are highly structured with course leaders following a clearly defined protocol. The predominant interventions are the generic Chronic Disease Self-Management Program (CDSMP) [16] and the disease-specific Arthritis Self-Management Course [17]. The difference between the two interventions is that the former is aimed at a broader audience with a wide range of different chronic conditions [18] as it is built on the assumption that people with any type of chronic disease face similar problems regarding the management of their condition [19].

In view of the large number of organizations offering Stanford programs, thereby investing a considerable amount of public and private funds, there is an urgent need to understand and document the true impact of this intervention. As none of the reviewed meta-analyses or systematic reviews focused on this specific protocol, but instead included a range of different self-management programs that not only differed across but also even within the same chronic disease [5], it was deemed necessary to prepare a systematic review of self-management trials that were largely following the Stanford curricula [16, 17]. Apart from a summary of the effectiveness of these interventions, the main objective of this research is to provide an in-depth investigation of the types of outcomes that studies are based on. While the present research summarizes the systematic review, our companion paper explores the pattern of effect sizes across studies and relates this pattern to the different types of outcome measures used.

Methods

Search strategy

Both systematic search and literature review were carried out in January 2007. The search was performed across the databases MEDLINE, EMBASE, CINAHL, and PsycINFO as recommended for systematic reviews [20].

The criteria and rationale for selecting studies for the systematic review were as follows:

1. Inclusion of studies evaluating disease-specific or generic self-management interventions comparable with the Stanford curricula. If studies did not directly refer to Stanford [16, 17], studies were selected that evaluated interventions that included at least two of the three keywords 'problem-solving', 'action planning', and 'relaxation'. To be included in the review, four characteristics had to be met by the self-management program:

    a. Interventions were delivered in a group setting;
    b. Interventions were based on a formal syllabus;
    c. Interventions ran between four and ten sessions within a period of 3 months;
    d. Interventions did not include any additional components such as exercise lessons, reinforcement techniques, individual consultations, and/or home visits.

2. Inclusion of studies between 1982 and 2006 as the first Stanford program was published in 1982 [21].
3. The search was limited to randomized controlled trials (RCTs), following the hierarchy of research designs where RCTs are considered the 'gold standard' [22].
4. Exclusion of studies that did not have sufficient power to detect a large-sized difference between intervention and control group means; that is, at $\alpha = 0.05$, a minimum sample size of $n = 26$ is required to detect a large difference between the means of two independent samples [23, 24].
5. Exclusion of studies that did not provide sufficient information on the outcome variables for calculating effect sizes (ES), and missing data could not be obtained from the authors.
6. Exclusion of studies that did not assess any self-report outcomes; that is, studies assessing outcomes such as cost-effectiveness or drug adherence were excluded.

7. Exclusion of studies on interventions for children or adolescents.

The search terms used in the systematic review were derived from other reviews in this area [5, 7] and were chosen in a way that the first part of the search string referred to the type of intervention, the second part referred to the study design, and the third part referred to the target group of the intervention. Hence, the search terms were ('patient education' or 'self-management') and ('randomised' or 'randomized' or 'RCT') and ('arthritis' or 'asthma' or 'chronic condition' or 'chronic disease' or 'chronic obstructive pulmonary disease' or 'congestive heart failure' or 'COPD' or 'diabetes' or 'fibromyalgia' or 'hypertension' or 'musculoskeletal' or 'osteoarthritis' or 'osteoporosis' or 'pain' or 'rheumatoid' or 'stress'). The following limits were set: in PsycINFO, the search was restricted to 'Journal articles only', languages 'English, German, Spanish' and 'Age groups 18 years and older'; in CINAHL, the search was restricted to above languages and 'all adult'. Additional studies were retrieved from the Cochrane Database of Systematic Reviews [25] and from reference lists of other reviews and meta-analyses.

Data analysis and presentation

For comparison across trials, ES were calculated for each outcome. Between-group treatment effects were calculated using Cohen's d [26], a commonly reported effect size index [27]. The 95 % confidence interval was estimated using standard errors obtained by multiplying the square root of the standard errors by 1.96 [28]. Additionally, within-group ES were calculated for each treatment condition separately. This was used to identify the source of between-group differences; for example, whether a large effect was caused by an improvement in the intervention group (IG) and/or a decline in the control group (CG) or whether a negligible between-group ES was caused by no change in either group or by a parallel change in both IG and CG subjects. Reported effects are presented such that positive ES reflect improvement and negative ES reflect decline. Results were interpreted as small (ES $\sim$ 0.2), medium (ES $\sim$ 0.5), or large (ES $\sim$ 0.8) effects [26].

While several studies included repeated measures with varying time points, only the first post-intervention assessment was used for the calculation of ES. Results are presented in a way that both the minimum and the maximum ES per outcome across studies are shown. Further, the median ES of the included studies were calculated. Hence, no summary scores such as those used in meta-analyses are reported as the aim of the review was to show and discuss the full range of effects across studies.

Results

The results of the systematic search are presented in Fig. 1. A total of 2,175 papers were identified in this search. After pre-screening the titles, 1,676 publications were excluded as they failed to meet all inclusion criteria. The majority of studies were rejected because they evaluated other types of interventions ($n = 861$), were a description of a program ($n = 319$), or children/adolescents were included ($n = 177$). Of the remaining 499 publications, all abstracts were screened. The majority of these again did not meet all inclusion criteria. This left 78 papers which were examined in full. Of these, 55 studies were excluded with the majority again describing different types of interventions such as classic patient education. Eight of the remaining 23 studies did not report sufficient data for the calculation of effect sizes. After contacting the authors, three researchers were able to provide most missing data [29–31], while no further information could be obtained from five authors [32–36]. As a result, the review is based on 18 studies.

The majority of the 18 trials investigated the effectiveness of arthritis-specific interventions. Seven evaluated the Stanford Arthritis Self-Management Course for osteoarthritis [37–39] or other musculoskeletal disorders [40–43]. Three trials evaluated alternative arthritis-specific interventions, with one being the 'Bone Up On Arthritis' program [44], while two explored interventions for rheumatoid arthritis [31, 45]. The remaining studies evaluated disease-specific interventions for people with back pain [46], chronic pain [47], and fibromyalgia [48], and four focused on the generic CDSMP [19, 30, 49, 50]. One of above trials compared lay-led with professional-led interventions, with each trial including a separate control group [41]. As we did not differentiate between 'modes of instruction' in the present research, the two trials were regarded as separate studies. Consequently, the total number of trials included in this review increased to 19.

Across trials, more than 70 different variables were assessed. As most of these were not assessed frequently enough to perform inter-study comparisons, this review is restricted to the most often reported outcomes. Hence, a total of 11 variables are shown in Table 1, with *pain*, *disability*, *depression*, and *self-efficacy* being the most commonly assessed outcomes. Further outcomes presented herein are *visits to physician*, *general health*, *fatigue*, *communication with physician*, *knowledge*, *anxiety*, and *physical functioning*. For a more detailed overview of the results, see Appendix.

The impact on *pain* was assessed across all but one study [45]. Reported ES varied greatly and ranged from some trials observing negative to negligible between-group effects [19, 29, 30, 40, 41, 44, 46, 47] to other studies showing small to medium positive between-group effects

Fig. 1 Flowchart of the search strategy
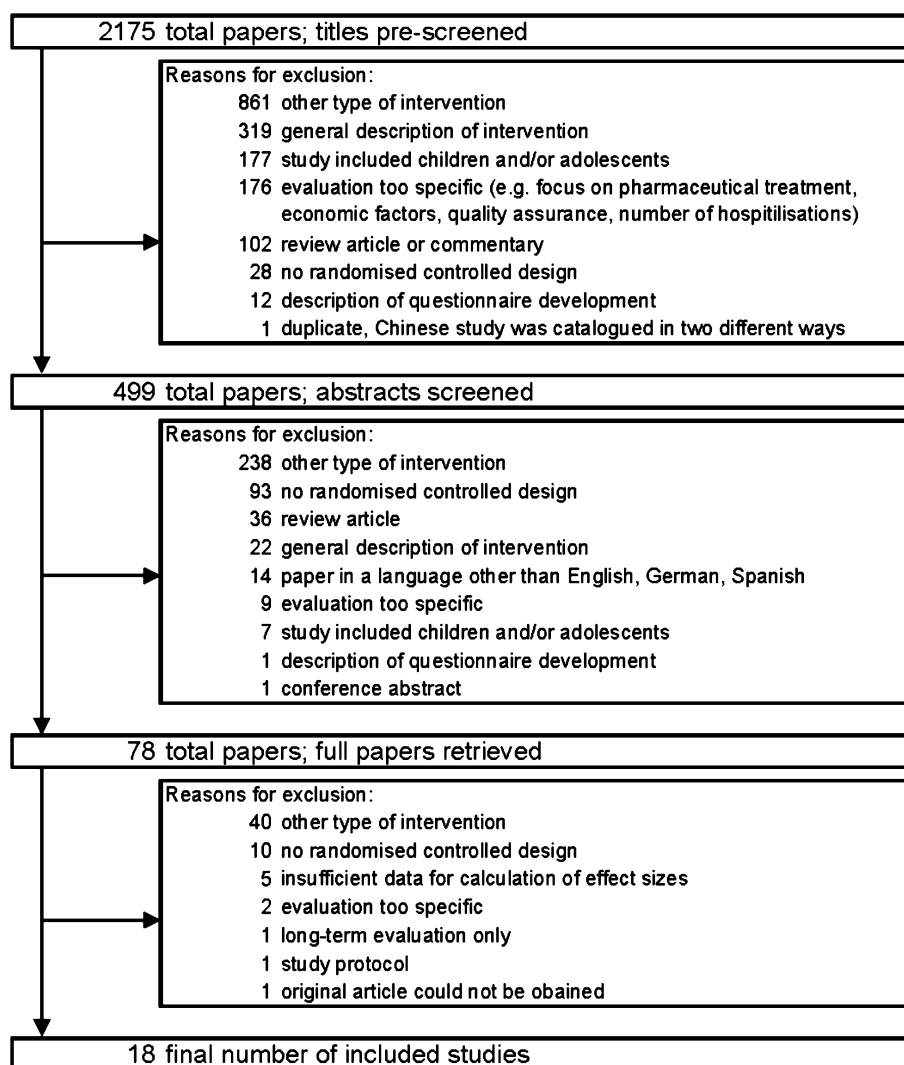
[31, 37–39, 42, 43, 48–50]. In three of these studies [38, 48, 49], however, between-group effects were largely influenced by increased pain in control group subjects.

Another outcome that was assessed frequently was *disability*. Reported effects again varied greatly across studies, with Cohen's d ranging from ES = −0.18 [39] to ES = 1.42 [45]. Half of included studies showed negligible to small effects [19, 41–44, 47], while four reported small to medium between-group effects [31, 46, 48, 49]. Within-group effects for intervention group subjects showed a similar range of benefits, with almost all effects ranging between ES = 0.0 and ES = 0.26 [19, 31, 41–44, 48, 49].

Depression was reported in 10 trials. Both between- and within-group effects varied greatly. While one trial showed medium effects [45] and two trials showed small between- and within-group effects [40, 48], most other studies reported negligible to small between-group ES [29–31, 42–44, 49]. In three of these studies, however, between-group effects were influenced by simultaneous improvements in both intervention and control group subjects [31, 42, 49].

Self-efficacy was also assessed in 10 trials. Reported results again varied greatly. Effects ranged from very small [37, 46] to above medium size between- and within-group effects [48]. Small to medium between- and within-group effects were reported in five studies [29, 30, 40, 49, 50]. Medium between-group ES were observed in another two trials [38, 42] of which one also showed medium within-group effects [42].

The number of *visits to physician* was assessed in eight studies. Calculated ES ranged from small decreases to small increases in the number of visits [19, 30, 40–42, 49, 50].

The impact of self-management courses on *general health* was assessed in seven trials. Similar to most previous outcomes, effects varied greatly between studies. They ranged from small negative between-group [46] to medium positive between- and within-group effects [50]. The median effect was small, with between-group effects of 0.16 and a within-group effect size of 0.17 for intervention group subjects.

Table 1 Effect sizes (ES) of most frequently assessed outcome measures in studies of chronic disease self-management programs largely following the Stanford curricula

Assessed outcome	Number of studies included	Minimum	Median	Maximum	Range
Pain[a]					
d	18	−0.28	0.10	0.43	0.71
IG		−0.04	0.20	0.75	0.79
CG		−0.34	0.11	0.80	1.14
Disability					
d	13	−0.18	0.14	1.42	1.60
IG		−0.20	0.15	1.28	1.48
CG		−0.43	−0.02	0.31	0.74
Depression[b]					
d	10	0.00	0.12	0.64	0.64
IG		0.07	0.22	0.57	0.50
CG		−0.06	0.04	0.36	0.42
Self-efficacy					
d	10	0.05	0.30	0.72	0.67
IG		0.02	0.40	0.64	0.62
CG		−0.34	0.01	0.27	0.61
Visits to physician					
d	8	−0.34	0.02	0.18	0.52
IG		−0.11	0.13	0.18	0.29
CG		0.02	0.12	0.21	0.19
General health					
d	7	−0.21	0.16	0.48	0.69
IG		−0.11	0.17	0.52	0.63
CG		−0.19	0.04	0.16	0.35
Fatigue					
d	6	−0.01	0.16	0.26	0.27
IG		0.12	0.17	0.40	0.28
CG		−0.04	0.05	0.13	0.17
Communication with physician					
d	5	−0.07	0.13	0.34	0.41
IG		0.04	0.23	0.49	0.45
CG		0.03	0.10	0.23	0.20
Knowledge					
d	5	−0.05	0.78	1.11	1.16
IG		0.37	0.95	1.28	0.91
CG		0.04	0.17	0.42	0.38
Anxiety					
d	4	0.02	0.12	0.20	0.18
IG		0.11	0.16	0.31	0.20
CG		0.00	0.08	0.13	0.13
Physical functioning					
d	4	−0.04	0.11	0.23	0.27
IG		−0.06	0.06	0.26	0.32
CG		−0.12	−0.01	0.03	0.15

CG effect size, control group, *IG* effect size, intervention group, *d* Cohen's d, group difference [ES (intervention group) − ES (control group)]

[a] Heuts et al. [37] reported knee and hip pain separately; these data are included as an average score of the two effect sizes

[b] Scholten et al. [45] assessed depression with the Freiburg Questionnaire of Coping with Illness (FQCI) and with the Beck Depression Index. These data are also presented as an average score

Six studies assessed *fatigue*. The majority of trials reported negligible to small between- and within-group ES [19, 29, 30, 40, 49]. The only exception was a Hispanic chronic disease self-management intervention with between-group effects of ES = 0.26 and within-group effects of ES = 0.40 for intervention group subjects [50].

Communication with the physician was assessed in five trials. Between- and within-group effects were negligible to

small [19, 30, 40, 49] in most studies. Larger effects were observed in a Spanish-speaking population with between-group effects of ES = 0.34 and within-group effects of ES = 0.49 [50].

Knowledge was also assessed in five of the included studies all of which were arthritis-specific interventions. In contrast to all other outcomes, both between- and within-group effects were generally medium or large [38, 41, 43, 44]. The only exception was the trial comparing lay-led courses with interventions run by health professionals. While the former showed medium-sized improvements in both intervention and control groups, the latter had a much larger impact on intervention group subjects [41].

The impact of self-management interventions on levels of *anxiety* was assessed in four trials. Overall, negligible to small effects were found, with within-group ES being consistently larger than between-group ES [29, 30, 40]. Largest within-group effects were reported by Taal et al. [31], showing an effect size of 0.31 for intervention group subjects compared with 0.13 for control group subjects [31].

Finally, *physical functioning* was also assessed in four trials. In contrast to most previously presented outcomes, between- and within-group effects were largely consistent, with all studies showing negligible [37, 40] or small effects [29, 48].

Discussion

This review summarizes the effectiveness of group-based chronic disease self-management courses. In contrast to other reviews in this area, the inclusion criteria were rather strict in that only studies were included that either followed or were reasonably similar to the Stanford curricula [16, 17]. As the majority of included trials described programs that were targeted at people with musculoskeletal conditions, this review can best be compared with reviews on arthritis. Although our review is restricted to Stanford-type programs and only median scores are presented herein, reported effects are largely similar to summary scores of reviews on arthritis. For example, outcomes *depression* [10], *disability* [6, 10, 12], and *pain* [6, 7, 10, 12] showed negligible to small ES across all publications. The only deviation was found for *self-efficacy*, with the present review showing somewhat larger effects than reported by others [6]. In sum, people with arthritis seem to only marginally benefit from participating in self-management programs across a range of outcomes [5–8, 11, 12].

Above findings are discouraging and seem to be in stark contrast to our previous research where we showed that on average one-third of participants received substantial benefits from attending such courses. That is, expressed in terms of number needed to treat (NNT), three persons had to attend a chronic disease self-management course for one person to receive substantial benefits [14] which is a rather good result. However, several aspects need to be considered when interpreting results presented in this review. First, it needs to be discussed whether changes are to be expected in some outcomes that are frequently assessed. For example, psycho-educational programs such as self-management courses are neither aimed at nor can they achieve a reduction in levels of *pain*. In contrast, these programs are aimed at providing individuals with the skills on how to cope with symptoms and manage episodes. That is, the aim of self-management education is the reduction in the perception of pain, that is, the effect that the perceived pain has on the patient, rather than a reduction in the actual level of pain [5]. Hence, if such outcomes are included, it is important to define the intent of the outcome measure, that is, level of pain versus pain coping and self-management skills.

Second, closely related to the previous aspect, it is questionable whether studies assess the most pertinent outcomes that self-management programs are supposed to impact on. As was observed in a narrative review [5], studies frequently assess outcomes that are not particularly targeted by these interventions. While the Stanford curricula include topics on communication with the physician, emotions and self-efficacy, it is questionable whether self-management programs are able to have an impact on variables such as disability, fatigue, physical functioning, and again pain as these outcomes are not specifically targeted by the programs' curricula. An overview of intended impacts of chronic disease self-management programs, incorporating both the patients' and the health professionals' perspectives, has been provided through the Health Education Impact Questionnaire (heiQ). This questionnaire was developed during a comprehensive stakeholder consultation phase and can be used to reliably evaluate chronic disease self-management programs [14, 51, 52]. Instead of focusing on physiological or clinical outcomes, the heiQ outcomes cover areas such as behaviors, skills, attitudes, self-monitoring, health services navigation, and emotional distress [52].

Third, a further challenge in interpreting outcomes of self-management programs concerns the time frame in which outcomes can be expected to occur. The trials of the present review assessed post-intervention outcomes ranging from direct post-course assessment to several months after the intervention. It is beyond the scope of the present research to consider this dimension further; however, studies that are concerned with the effectiveness of chronic disease self-management education should take the dimension 'time' into account as the different types of outcomes can be expected to occur at different points in

time. A program logic model of impacts of self-management education can serve as a guide to categorize outcomes into short-, medium-, and long-term effects [53].

Finally, when comparing program outcomes for people with arthritis with other chronic illnesses, it becomes clear that studies rely on a different set of outcomes across the different types of interventions. For example, outcomes for participants with diabetes or hypertension can be both self-report and clinical. In contrast, there are no objective biological measures of disease severity in musculoskeletal diseases. As a result, when looking at program effectiveness, evaluators have to exclusively rely on participant self-report [5]. In view of large differences in accuracy of measuring self-report variables compared with clinically assessed outcomes, it is plausible that evaluations do not accurately reflect program effects. This hypothesis is supported by the Quality of Life Appraisal model by Schwartz and Rapkin [54] where 'appraisal' denotes the cognitive process carried out when responding to a question. The model describes an increase in measurement error and bias with increasing complexity involved in the response process [54]. In our companion paper [reference to potential companion paper], we describe a practical application of the model where program outcomes presented herein are allocated to the three categories performance-, perception-, and evaluation-based measures as introduced by Schwartz and Rapkin [54].

The present research has some limitations. Even though this review was defined narrowly, with only Stanford-type interventions included, comparability of included trials is still difficult; that is, over 70 variables were assessed across studies, and data were often collected with different measurement instruments. The former aspect clearly shows that, while only comparable interventions were included where according program objectives should be similar if not identical, it is surprising that different research groups

expect impacts across such large range of outcomes. Further, the aforementioned range of instruments is of concern. While this aspect was not specifically considered in this research, a careful review of questionnaires suitable for the evaluation of self-management programs should be carried out as the different instruments differ in their relative sensitivity to measure change [5], and it seems inappropriate to merge scores from different tests. To make results comparable across studies—while taking into account the objectives of different self-management education interventions—a standardized suite of outcome variables and according instruments needs to be defined to evaluate self-management programs and make results comparable across trials. For example, the heiQ [52] may well be suitable given its robust psychometrics and wide application, including its function as the national quality and monitoring tool for the UK Expert Patients Program [55].

In summary, this research confirms other reviews in this area, suggesting that the Stanford CDSMP and other programs have nil to small effects on program participants. However, it also suggests that, when looking at the types of outcomes that trials are based on, alternative explanations for these results seem probable. As evaluations heavily rely on participant self-report, that is, outcomes that show negligible to small effects, current approaches to program evaluation may not be sufficient to assess the true impact of chronic disease self-management education. An in-depth exploration of the types of outcomes that the reviewed trials are based on is described in a separate paper.

Appendix

See Table 2.

Table 2 Included studies, types of interventions, outcome measures, and effects (Cohen's d [95 % confidence interval, CI], effect sizes for both intervention group [IG] and control group [CG])

Study/intervention	Outcome measures	Effect size[a] Cohen's d	95 % CI Lower	Upper	Effect size separately[b] IG	CG
Lorig et al. [41] (ASMP, led by health professional)	4 Months				n = 29	n = 29
	Disability	0.00	−0.51	0.51	0.14	0.14
	Knowledge	0.62	0.09	1.14	1.00	0.42
	Pain	−0.22	−0.73	0.30	0.14	0.38
	Visits to physician	−0.34	−0.86	0.18	−0.11	0.21
Lorig et al. [41] (ASMP, led by peers)	4 Months				n = 27	n = 29
	Disability	0.13	−0.39	0.66	0.25	0.14
	Knowledge	−0.05	−0.58	0.47	0.37	0.42
	Pain	−0.28	−0.81	0.24	0.07	0.38
	Visits to physician	−0.10	−0.62	0.42	0.07	0.21

Table 2 continued

Study/intervention	Outcome measures	Effect size[a] Cohen's d	95 % CI Lower	Upper	Effect size separately[b] IG	CG
Lorig et al. [43] (ASMP)	4 Months				$n = 501$	$n = 206$
	Depression	0.11	−0.06	0.27	0.10	−0.01
	Disability	0.05	−0.11	0.21	0.02	−0.03
	Knowledge	0.78	0.62	0.95	0.83	0.04
	Pain	0.15	−0.01	0.31	0.23	0.07
Goeppinger et al. [44][c] (BUOA)	4 Months				$n = 100$	$n = 153$
	Depression	0.12	−0.13	0.37	0.12	0.00
	Disability	0.02	−0.23	0.27	0.00	−0.02
	Knowledge	0.85	0.59	1.11	0.95	0.10
	Pain	0.08	−0.17	0.34	0.21	0.13
Keefe et al. [39] (ASMP)	Post-treatment				$n = 36$	$n = 31$
	Disability	−0.18	−0.66	0.30	−0.20	−0.03
	Pain	−0.23	−0.71	0.25	−0.04	0.20
Taal et al. [31] (RA course, based on social learning and ASMP)	Post-treatment				$n = 27$	$n = 30$
	Anxiety	0.20	−0.32	0.72	0.31	0.13
	Depression	0.05	−0.47	0.57	0.35	0.36
	Disability	0.42	−0.10	0.95	0.02	−0.43
	Pain	0.18	−0.34	0.70	0.26	0.07
Burckhardt et al. [29] (Fibromyalgia self-management course)	6 weeks				$n = 28$	$n = 30$
	Anxiety	0.10	−0.41	0.62	0.11	0.00
	Depression	0.00	−0.52	0.52	0.12	0.11
	Fatigue	0.14	−0.37	0.66	0.19	0.08
	Pain	−0.14	−0.65	0.38	0.07	0.21
	Pain, tender points	0.29	−0.23	0.81	0.31	0.04
	Physical function	0.23	−0.29	0.74	0.26	0.00
	Self-efficacy, function	0.18	−0.33	0.70	0.13	−0.05
	Self-efficacy, other	0.32	−0.20	0.84	0.44	0.07
	Self-efficacy, pain	0.08	−0.43	0.60	0.13	0.04
LeFort et al. [48] (modified ASMP for chronic pain)	6 weeks				$n = 52$	$n = 50$
	Depression	0.22	−0.17	0.61	0.17	−0.04
	Disability	0.30	−0.09	0.69	0.26	−0.03
	General health	0.16	−0.23	0.55	0.17	0.00
	Pain quality	0.43	0.04	0.83	0.21	−0.23
	Physical functioning	0.14	−0.25	0.52	0.12	−0.01
	Self-efficacy	0.72	0.32	1.12	0.64	−0.11
Von Korff et al. [47] (back pain, after Stanford)	3 Months				$n = 124$	$n = 121$
	Disability	0.15	−0.10	0.40	0.48	0.31
	Pain	−0.07	−0.32	0.18	0.75	0.80
Lorig et al. [19] (Spanish ASMP)	4 Months				$n = 189$	$n = 97$
	Depression	0.06	−0.19	0.30	0.26	0.20
	Disability	0.14	−0.10	0.39	0.15	0.00
	General health	0.24	−0.01	0.48	0.41	0.12
	Pain	0.37	0.12	0.61	0.35	−0.01
	Self-efficacy	0.48	0.24	0.73	0.48	−0.02
	Visits to physician	−0.08	−0.32	0.17	0.02	0.13

Table 2 continued

Study/intervention	Outcome measures	Effect size[a]	95 % CI		Effect size separately[b]	
		Cohen's d	Lower	Upper	IG	CG
Lorig et al. [16] (CDSMP)	6 Months				$n = 561$	$n = 391$
	Communication with physician	0.13	0.00	0.25	0.22	0.09
	Disability	0.08	−0.05	0.21	0.03	−0.05
	Energy/fatigue	0.11	−0.02	0.24	0.13	0.02
	Pain	0.02	−0.11	0.15	0.12	0.09
	Self-rated health	0.12	−0.01	0.25	0.10	−0.02
	Visits to physician	0.04	−0.09	0.17	0.14	0.09
Scholten et al. [45] (multidisciplinary arthritis training, based on Stanford)	Post-treatment				$n = 38$	$n = 30$
	Depression	0.42	−0.07	0.90	0.30	−0.09
	Depression, Beck	0.86	0.36	1.36	0.84	−0.03
	Disability	1.42	0.88	1.95	1.28	0.00
Barlow et al. [40] (ASMP)	4 Months				$n = 234$	$n = 189$
	Anxiety	0.14	−0.05	0.33	0.21	0.08
	Communication with physician	0.19	0.00	0.38	0.23	0.03
	Depression	0.25	0.06	0.44	0.27	0.04
	Fatigue	0.18	−0.02	0.37	0.17	−0.02
	Pain	0.03	−0.16	0.22	0.12	0.09
	Physical functioning	−0.04	−0.23	0.15	−0.01	0.03
	Self-efficacy, symptoms	0.31	0.12	0.50	0.43	0.13
	Self-efficacy, pain	0.26	0.07	0.45	0.41	0.14
					$n = 86$	$n = 78$
	Physician visits: other	−0.36	−0.67	−0.05	−0.17	0.19
	Visits to physician to discuss arthritis	0.04	−0.27	0.35	0.15	0.08
Hopman-Rock et al. [38] (ASMP, modified for osteoarthritis)	Post-treatment				$n = 56$	$n = 49$
	Knowledge	1.11	0.70	1.52	1.28	0.17
	Pain	0.43	0.04	0.81	0.10	−0.34
	Self-efficacy	0.51	0.12	0.89	0.19	−0.34
Fu et al. [49] (CDSMP modified for Chinese culture)	6 Months				$n = 430$	$n = 349$
	Communication with physician	−0.07	−0.21	0.07	0.04	0.10
	Depression	0.13	−0.01	0.27	0.28	0.16
	Disability	0.32	0.18	0.46	0.25	−0.05
	Fatigue	0.20	0.06	0.34	0.16	−0.04
	Pain	0.19	0.05	0.33	0.02	−0.17
	Self-efficacy, symptoms	0.31	0.17	0.45	0.23	−0.08
	Self-efficacy, disease	0.29	0.15	0.43	0.10	−0.19
	Self-rated health	0.42	0.28	0.56	0.47	0.05
	Visits to physician	0.02	−0.12	0.16	0.11	0.11
Lorig et al. [50] (Spanish CDSMP)	4 Months				$n = 265$	$n = 178$
	Communication with physician	0.34	0.14	0.53	0.49	0.16
	Fatigue	0.26	0.07	0.46	0.40	0.11
	Pain	0.23	0.04	0.42	0.36	0.13
	Self-efficacy	0.16	−0.03	0.35	0.44	0.27
	Self-rated health	0.48	0.29	0.68	0.52	0.04
	Visits to physician	0.18	−0.01	0.37	0.18	0.02

Table 2 continued

Study/intervention	Outcome measures	Effect size[a]	95 % CI		Effect size separately[b]	
		Cohen's d	Lower	Upper	IG	CG
Griffiths et al. [30] (CDSMP in a South Asian group)	4 Months				$n = 221$	$n = 218$
	Anxiety	0.02	−0.16	0.21	0.11	0.08
	Communication with physician	0.09	−0.10	0.27	0.32	0.23
	Depression	0.04	−0.15	0.22	0.07	0.04
	Fatigue	−0.01	−0.20	0.18	0.12	0.13
	Pain	0.03	−0.16	0.21	0.24	0.22
	Self-efficacy	0.13	−0.06	0.32	0.36	0.24
	Visits to physician	0.01	−0.18	0.20	0.17	0.15
Haas et al. [46] (CDSMP low back pain)	6 Months				$n = 60$	$n = 49$
	Disability	0.26	−0.12	0.64	0.41	0.13
	General health	−0.21	−0.59	0.17	−0.05	0.16
	Pain	0.03	−0.34	0.41	0.27	0.23
	Self-efficacy, symptoms	0.12	−0.25	0.50	0.02	−0.11
	Self-efficacy, pain	−0.04	−0.42	0.34	0.02	0.06
Heuts et al. [37] (ASMP modified for osteoarthritis)	3 Months				$n = 132$	$n = 141$
	General health	0.09	−0.15	0.33	−0.11	−0.19
	Pain, hip	−0.02	−0.26	0.22	0.08	0.10
	Pain, knee	0.25	0.01	0.49	0.28	0.00
	Physical functioning	0.08	−0.16	0.31	−0.06	−0.12
	Self-efficacy	0.05	−0.18	0.29	0.10	0.04

ASMP arthritis self-management program, *BUOA* bone up on arthritis, *CDSMP* chronic disease self-management program, *CG* control group, *IG* intervention group, *RA* rheumatoid arthritis

[a] Effect sizes were based on Cohen's d [26]

[b] Effect sizes were calculated separately for the intervention and the control group. Effect sizes refer to the follow-up score at time X minus baseline score divided by the SD at baseline of the respective group

[c] Goeppinger et al. [44] only reported pooled SD of respective pre-test scores. Hence, where effect sizes are reported separately, each group's change score is divided by this pooled SD. This differs from the other studies in which the individual change scores are divided by each group's individual SD

References

1. Yach, D., Hawkes, C., Gould, C. L., & Hofman, K. J. (2004). The global burden of chronic diseases—overcoming impediments to prevention and control. *JAMA, 291*, 2616–2622.

2. Abegunde, D. O., Mathers, C. D., Adam, T., Ortegon, M., & Strong, K. (2007). The burden and costs of chronic diseases in low-income and middle-income countries. *Lancet, 370*(9603), 1929–1938.

3. Strong, K., Mathers, C., Leeder, S., & Beaglehole, R. (2005). Chronic diseases 1—Preventing chronic diseases: How many lives can we save? *Lancet, 366*, 1578–1582.

4. Ofman, J. J., Badamgarav, E., Henning, J. M., Knight, K., Gano, A. D., Jr., Levan, R. K., et al. (2004). Does disease management improve clinical and economic outcomes in patients with chronic diseases? A systematic review. *American Journal of Medicine, 117*, 182–192.

5. Newman, S., Steed, L., & Mulligan, K. (2004). Self-management interventions for chronic illness. *Lancet, 364*(9444), 1523–1537.

6. Astin, J. A., Beckner, W., Soeken, K., Hochberg, M. C., & Berman, B. (2002). Psychological interventions for rheumatoid arthritis: A meta-analysis of randomized controlled trials. *Arthritis and Rheumatism, 47*(3), 291–302.

7. Chodosh, J., Morton, S. C., Mojica, W., Maglione, M., Suttorp, M. J., Hilton, L., et al. (2005). Meta-analysis: Chronic disease self-management programs for older adults. *Annals of Internal Medicine, 143*(6), 427–438.

8. Devos-Comby, L., Cronan, T., & Roesch, S. C. (2006). Do exercise and self-management interventions benefit patients with osteoarthritis of the knee? A metaanalytic review. *Journal of Rheumatology, 33*(4), 744–756.

9. Ismail, K., Winkley, K., & Rabe-Hesketh, S. (2004). Systematic review and meta-analysis of randomised controlled trials of psychological interventions to improve glycaemic control in patients with type 2 diabetes. *Lancet, 363*(9421), 1589–1597.

10. Mullen, P. D., Laville, E. A., Biddle, A. K., & Lorig, K. (1987). Efficacy of psychoeducational interventions on pain, depression, and disability in people with arthritis: A meta-analysis. *Journal of Rheumatology, 14*(Suppl 15), 33–39.

11. Warsi, A., Wang, P. S., LaValley, M. P., Avorn, J., & Solomon, D. H. (2004). Self-management education programs in chronic disease: A systematic review and methodological critique of the literature. *Archives of Internal Medicine, 164*(15), 1641–1649.

12. Warsi, A., LaValley, M. P., Wang, P. S., Avorn, J., & Solomon, D. H. (2003). Arthritis self-management education programs: A

meta-analysis of the effect on pain and disability. *Arthritis and Rheumatism, 48*(8), 2207–2213.

13. Jordan, J. E., & Osborne, R. H. (2007). Chronic disease self-management education programs: Challenges ahead. *MJA, 186*(2), 84–87.

14. Nolte, S., Elsworth, G. R., Sinclair, A. J., & Osborne, R. H. (2007). The extent and breadth of benefits from participating in chronic disease self-management courses: A national patient-reported outcomes survey. *Patient Education and Counseling, 65*(3), 351–360.

15. Stanford Patient Education Research Center. (2012). *Organizations Licensed to Offer the Chronic Disease Self-Management Program*. Retrieved 04.09.2012, from http://patienteducation.stanford.edu/organ/cdsites.html#aus.

16. Lorig, K. R., González, V. M., & Laurent, D. D. (1999). *The Chronic Disease Self-Management Program: Leaders Manual*. Palo Alto: Stanford University.

17. Lorig, K., Lubeck, D., Kraines, R. G., Seleznick, M., & Holman, H. R. (1985). Outcomes of self-help education for patients with arthritis. *Arthritis and Rheumatism, 28*(6), 680–685.

18. Ritter, P. L., Lee, J., & Lorig, K. (2011). Moderators of chronic disease self-management programs: Who benefits? *Chronic Illn, 7*(2), 162–172.

19. Lorig, K. R., Sobel, D. S., Stewart, A. L., Brown, B. W., Bandura, A., Ritter, P., et al. (1999). Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization. *Medical Care, 37*(1), 5–14.

20. Taal, E., Riemsma, R. P., Kirwan, J. R., & Rasker, J. J. (2004). What are the real effects of arthritis self-management education programs on pain and disability? Comment on the article by Warsi et al. *Arthritis and Rheumatism, 50*(3), 1012–1013.

21. Lorig, K. R. (1982). Arthritis self-management: A patient education program. *Rehabilitation Nursing, 8*, 16–20.

22. Sackett, D. L. (1994). The cochrane collaboration. *ACP Journal Club, 120*(Suppl 3), A-11.

23. Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

24. Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A., & Fayers, P. M. (2006). The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research, 15*(9), 1533–1550.

25. The Cochrane Collaboration. (2007). Cochrane reviews. Retrieved 03.01.2007, from http://www.cochrane.org/reviews/.

26. Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

27. Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods, 1*, 331–340.

28. Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

29. Burckhardt, C. S., Mannerkorpi, K., Hedenberg, L., & Bjelle, A. (1994). A randomized, controlled clinical trial of education and physical training for women with fibromyalgia. *Journal of Rheumatology, 21*, 714–720.

30. Griffiths, C., Motlib, J., Azad, A., Ramsay, J., Eldridge, S., Feder, G., et al. (2005). Randomised controlled trial of a lay-led self-management programme for Bangladeshi patients with chronic disease. *British Journal of General Practice, 55*, 831–837.

31. Taal, E., Riemsma, R. P., Brus, H. L., Seydel, E. R., Rasker, J. J., & Wiegman, O. (1993). Group education for patients with rheumatoid arthritis. *Patient Education and Counseling, 20*(2–3), 177–187.

32. Clark, N. M., Janz, N. K., Becker, M. H., Schork, M. A., Wheeler, J., Liang, J., et al. (1992). Impact of self-management education on the functional health status of older adults with heart disease. *Gerontologist, 32*, 438–443.

33. Cohen, J. L., van Houten Sauter, S., DeVellis, R. F., & McEvoy DeVellis, B. (1986). Evaluation of arthritis self-management courses led by laypersons and by professionals. *Arthritis and Rheumatism, 29*(3), 388–393.

34. Solomon, D. H., Warsi, A., Brown-Stevenson, T., Farrell, M., Gauthier, S., Mikels, D., et al. (2002). Does self-management education benefit all populations with arthritis? A randomized controlled trial in a primary care physician network. *Journal of Rheumatology, 29*(2), 362–368.

35. Swerissen, H., Belfrage, J., Weeks, A., Jordan, L., Walker, C., Furler, J., et al. (2006). A randomised control trial of a self-management program for people with a chronic illness from Vietnamese, Chinese, Italian and Greek backgrounds. *Patient Education and Counseling, 64*(1–3), 360–368.

36. Worth, H. (2002). Effects of patient education in asthma and COPD—What is provable? *Medizinische Klinik, 97*(2, Suppl), 20–24.

37. Heuts, P. H. T. G., de Bie, R., Drietelaar, M., Aretz, K., Hopman-Rock, M., Bastiaenen, C. H. G., et al. (2005). Self-management in osteoarthritis of hip or knee: A randomized clinical trial in a primary healthcare setting. *Journal of Rheumatology, 32*(3), 543–549.

38. Hopman-Rock, M., & Westhoff, M. H. (2000). The effects of a health educational and exercise program for older adults with osteoarthritis of the hip or knee. *Journal of Rheumatology, 27*(8), 1947–1954.

39. Keefe, F. J., Caldwell, D. S., Wiliams, D. A., Gil, K. M., Mitchell, D., Robertson, C., et al. (1990). Pain coping skills training in the management of osteoarthritic knee pain: A comparative study. *Behavior Therapy, 21*, 49–62.

40. Barlow, J. H., Turner, A. P., & Wright, C. C. (2000). A randomized controlled study of the Arthritis Self-Management Programme in the UK. *Health Education Research, 15*(6), 665–680.

41. Lorig, K., Feigenbaum, P., Regan, C., Ung, E., Chastain, R. L., & Holman, H. R. (1986). A comparison of lay-taught and professional-taught arthritis self-management courses. *Journal of Rheumatology, 13*(4), 763–767.

42. Lorig, K., González, V. M., & Ritter, P. (1999). Community-based Spanish language arthritis education program: A randomized trial. *Medical Care, 37*(9), 957–963.

43. Lorig, K., Seleznick, M., Lubeck, D., Ung, E., Chastain, R. L., & Holman, H. R. (1989). The beneficial outcomes of the arthritis self-management course are not adequately explained by behavior change. *Arthritis and Rheumatism, 32*(1), 91–95.

44. Goeppinger, J., Arthur, M. W., Baglioni, A. J., Jr., Brunk, S. E., & Brunner, C. M. (1989). A reexamination of the effectiveness of self-care education for persons with arthritis. *Arthritis and Rheumatism, 32*(6), 706–716.

45. Scholten, C., Brodowicz, T., Graninger, W., Gardavsky, I., Pils, K., Pesau, B., et al. (1999). Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. *Archives of Physical Medicine and Rehabilitation, 80*, 1282–1287.

46. Haas, M., Groupp, E., Muench, J., Kraemer, D., Brummel-Smith, K., Sharma, R., et al. (2005). Chronic disease self-management program for low back pain in the elderly. *Journal of Manipulative and Physiological Therapeutics, 28*(4), 228–237.

47. Von Korff, M., Moore, J. E., Lorig, K., Cherkin, D. C., Saunders, K., González, V. M., et al. (1998). A randomized trial of a lay person-led self-management group intervention for back pain patients in primary care. *Spine, 23*(23), 2608–2615.

48. LeFort, S. M., Gray-Donald, K., Rowat, K. M., & Jeans, M. E. (1998). Randomized controlled trial of a community-based psychoeducation program for the self-management of chronic pain. *Pain, 74*(2–3), 297–306.

49. Fu, D., Fu, H., McGowan, P., Shen, Y.-E., Zhu, L., Yang, H., et al. (2003). Implementation and quantitative evaluation of chronic disease self-management programme in Shanghai, China: Randomized controlled trial. *Bulletin of the World Health Organization, 81*(3), 174–182.

50. Lorig, K. R., Ritter, P. L., & González, V. M. (2003). Hispanic chronic disease self-management—A randomized community-based outcome trial. *Nursing Research, 52*(6), 361–369.

51. Nolte, S., Elsworth, G. R., Sinclair, A. J., & Osborne, R. H. (2009). Tests of measurement invariance failed to support the application of the "then-test". *Journal of Clinical Epidemiology, 62*(11), 1173–1180.

52. Osborne, R. H., Elsworth, G. R., & Whitfield, K. (2007). The Health Education Impact Questionnaire (HEIQ): An outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient Education and Counseling, 66*, 192–201.

53. Osborne, R. H., Spinks, J. M., & Wicks, I. P. (2004). Patient education and self-management programs in arthritis. *MJA, 180*, S23–S26.

54. Schwartz, C. E., & Rapkin, B. D. (2004). Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. Health qual life outcomes retrieved April 15, 2006, from http://www.hqlo.com/content/2/1/16.

55. Osborne, R. H., Batterham, R., & Livingston, J. (2011). The evaluation of chronic disease self-management support across settings: The international experience of the health education impact questionnaire monitoring system. *Nursing Clinics of North America, 46*(3), 255–270.