

OPEN

Early Psychological Interventions for Somatic Symptom Disorder and Functional Somatic Syndromes: A Systematic Review and Meta-Analysis

Lukas Berezowski, MSc, Lea Ludwig, PhD, Alexandra Martin, PhD,
Bernd Löwe, MD, and Meike C. Shedden-Mora, PhD

ABSTRACT

Objective: Psychological treatments for somatic symptom disorder and functional somatic syndromes (SSD/FSS) achieve moderate effects only, potentially because of the high chronicity in these patients. Therefore, we aimed to evaluate whether early treatment, that is, treatment in populations at risk or with recent onset, improves outcome.

Methods: We conducted a systematic review and meta-analysis of (cluster-)randomized controlled trials evaluating early psychological interventions in the prevention and treatment of SSD/FSS in adults compared with inactive control conditions, standard care, or placebo. Individuals at risk for SSD/FSS, suffering from subthreshold symptoms or new onsets of SSD/FSS, or presenting with SSD/FSS for the first time were included.

Results: We identified 30 eligible studies, mostly examining pain-related conditions. Interventions were diverse, ranging from bibliotherapy to cognitive-behavioral therapy. We found positive effects on depression post-treatment (Hedges' $g = 0.12$ [95% confidence interval = 0.03–0.2], $k = 5$) as well as on somatic symptom severity ($g = 0.25$ [0.096–0.41], $k = 17$) and health care utilization ($g = 0.31$ [0.18–0.44], $k = 3$) at follow-up. However, because of a high risk of bias, sensitivity to corrections for meta-bias, and missing outcome data, findings should be interpreted cautiously.

Conclusions: Our review shows that targeting SSD/FSS at an early stage represents a conceptual and practical challenge. Readily accessible interventions addressing transsyndromal processes of SSD/FSS development and consolidation are highly needed. Future studies are needed to evaluate individuals with diverse symptoms, examine symptom history thoroughly, use placebo controls, and report outcomes completely to determine the efficacy of early psychological interventions for SSD/FSS.

PROSPERO Registration: CRD42020140122.

Key words: somatic symptom disorder, functional somatic syndrome, bodily distress, somatoform disorder, early psychological intervention, prevention.

INTRODUCTION

Bodily complaints are common and may be burdensome for the affected individual. In many cases, bodily complaints occur without a well-defined biomedical cause or the individual's suffering is not sufficiently explained by structural damage or dysfunction (1). There are numerous concepts and diagnoses to describe such complaints, such as the *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) diagnoses *somatic symptom disorder* (SSD) and *conversion disorder (functional neurological symptom disorder)* (2). Concurrently, the umbrella term functional somatic syndrome (FSS) subsumes various syndromal diagnoses specific to different medical specialties, for example, irritable bowel syndrome or chronic fatigue syndrome (3). Of note, all these different concepts might capture different subtypes of a single,

underlying clinical phenotype (4–6). In the following, we will use the terms SSD/FSS to refer to any clinically relevant bodily complaints without a well-defined biomedical cause to use a terminology familiar both to a psychological and a medical audience.

SSD/FSS are highly prevalent in the general population (7,8), in primary care (9), and in patients of a wide range of medical specialties (6). Compared with individuals suffering from well-defined biomedical conditions, individuals suffering from SSD/FSS show at least the same, if not worse, outcomes (10,11). Moreover,

CBT = cognitive-behavioral therapy, **CI** = confidence interval, **DUI** = duration of untreated illness, **g** = Hedges g , **HrQoL** = health-related quality of life, **k** = number of studies, **RR** = risk ratio, **SC/TAU** = standard care/treatment as usual, **SSD/FSS** = somatic symptom disorder/functional somatic syndrome

SDC Supplemental Digital Content

From the Clinical Psychology and Psychotherapy, Institute of Psychology, Faculty of Psychology and Movement Sciences (Berezowski, Ludwig), University of Hamburg, Hamburg; Department of Clinical Psychology and Psychotherapy (Martin), University of Wuppertal, Wuppertal; Department of Psychosomatic Medicine and Psychotherapy (Löwe, Shedden-Mora), University Medical Center Hamburg-Eppendorf; and Department of Psychology (Shedden-Mora), University Medical Center Hamburg-Eppendorf & Medical School Hamburg, Hamburg, Germany.

Address correspondence to Meike C. Shedden-Mora, PhD, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany. E-mail: m.shedden-mora@uke.de

Received for publication March 17, 2021; revision received August 20, 2021.

DOI: 10.1097/PSY.0000000000001011

Copyright © 2021 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Psychosomatic Society. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

individuals suffering from SSD/FSS display an elevated use of health care services (12), potentially exposing them to iatrogenic harm, for example, by unnecessary investigations (13).

Currently, multidisciplinary treatment is recommended for SSD/FSS, with psychological treatment being a key component (1,14). Meta-analytic evidence supports the efficacy of psychological treatments for SSD/FSS (15–18). In these meta-analyses, however, between-group effect sizes of psychological treatments are generally only small to moderate, typically around Cohen's $d = 0.11$ to 0.40 .

One possible explanation for these small effects could be a long duration of untreated illness (DUI) in these patients. The DUI describes the time frame between the onset of a disorder and the initiation of adequate treatment (19) and has been linked to worse outcome (20). The mean symptom duration reported in current reviews evaluating psychological treatments for SSD/FSS ranged from 3 to 25 years (16–18,21). In primary care, Herzog et al. (19) reported a mean DUI of 25.2 years. Accordingly, offering interventions at an earlier time might improve clinical outcome.

The appropriate conceptualization of early interventions for SSD/FSS warrants special consideration for several reasons. Only a small proportion of patients exposed to a risk factor develops chronic symptoms, for example, after a gastrointestinal infection (22). Furthermore, first symptoms of SSD/FSS often develop in childhood and adolescence (23) and show substantial fluctuations over time (24–26), complicating the definition and assessment of illness duration. Even when fully manifested, SSD/FSS are severely underrecognized (e.g., (27)) because of numerous structural as well as patient- and physician-related factors (26). Consequently, both a clear-cut differentiation of early intervention types for SSD/FSS and their respective implementation are currently challenging.

The aim of the present study was to systematically examine the efficacy of early psychological interventions in preventing and treating SSD/FSS compared with inactive control treatments, standard care, or placebo in adults. We conducted a systematic review and meta-analysis of (cluster-)randomized controlled trials evaluating the effect of early psychological interventions on core outcome domains in SSD/FSS research (28).

The methodology of this review mirrors two goals which we wanted to achieve. First, we integrated various subtypes of SSD/FSS to comply with the current conceptualization of SSD/FSS as one clinical phenotype (3), additionally allowing us to ground our conclusions on a broader empirical basis. Second, we consider a rather inclusive definition of early psychological interventions more adequate to the nature of SSD/FSS. Beyond that, such a perspective allows us to identify promising avenues and shortcomings in past early intervention research more comprehensively. To the best of our knowledge, this is the first systematic review and meta-analysis on early psychological intervention approaches for SSD/FSS in general. We expect early psychological interventions to show positive effects. If effective, early psychological interventions could gain increasing importance in the prevention and management of SSD/FSS.

METHODS

Protocol and Registration

A protocol was developed before data collection (see Text in Section A, Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>). In this article, we describe the study in the way it was finally conducted. A detailed description of all amendments can be found in the Table in Section B,

Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>. This review was registered with PROSPERO before data collection (CRD42020140122).

Eligibility Criteria

Because the onset of an SSD/FSS cannot be clearly defined in many cases, we incorporated three different participant eligibility criteria to capture the whole population of interest for this review. Participants needed to be adults (18 years and older) fulfilling at least one of the following criteria: a) being at elevated risk for developing an SSD/FSS due to an acute event, for example, suffering from an acute gastroenteritis (22) (prevention population, “incident” definition); b) suffering from an SSD/FSS as diagnosed by a medical/mental health professional for a maximum of 12 months, or suffering from subthreshold functional symptoms, or exhibiting somatic symptoms without clear somatic etiology and indication for somatic treatment (early intervention population, “time” or “recent onset” definition); and c) first presentation with an SSD/FSS to health care provider (first presentation population, “help-seeking” definition).

Because the onset of an SSD/FSS cannot be clearly defined in many cases, we incorporated these three different participant eligibility criteria to capture the whole population of interest for this review, that is, individuals with acute symptoms, individuals in the process of developing persistent symptoms, and recent-onset SSD/FSS. The first criterion aims to capture populations with an elevated likelihood of developing an SSD/FSS due to an acute event, for example, suffering from an acute gastroenteritis or a whiplash trauma. Studies were included if the acute event was a known risk factor for developing SSD/FSS based on current evidence (e.g., (22,29)). The second criterion aims to capture populations with subthreshold symptoms or new onsets of the full picture of SSD/FSS. The third criterion aims to capture populations who seek professional help for their SSD/FSS for the first time, irrespective of the duration of symptoms. We incorporated this criterion because in routine care, the delivery of early psychological interventions is only possible if the affected individuals do seek help.

We looked for prospective (cluster-)randomized controlled trials evaluating the efficacy of psychological interventions in preventing or treating SSD/FSS compared with no treatment, wait-list controls, standard care/treatment as usual (SC/TAU), or placebo. The results of these studies had to be published between January 1, 1994, and April 30, 2020, in English or German language. Clinician-directed interventions were included if they aimed at fostering the use of psychological interventions and patient-level outcomes were reported. If psychological interventions were delivered in combination with other interventions, the psychological interventions had to take up the majority of the sessions.

Outcomes of interest in the meta-analyses were selected based on recommendations for intervention research in SSD/FSS (28). Primary outcomes were somatic symptom severity (self-report) and health-related quality of life (HrQoL; self-report). Secondary outcomes were unwanted negative treatment effects, diagnostic status concerning SSD/FSS (clinician-rated), anxiety, depression, health care utilization (i.e., number of doctor visits), and consumer satisfaction. Studies did not need to measure one of our outcomes of interest to be included in the review but were excluded from the meta-analyses, accordingly.

Search

Searches were run between September 9 and September 11, 2019, as well as between May 4 and May 5, 2020, in PubMed (NCBI), PsycINFO (Ovid), and Web of Science (Clarivate Analytics). We developed a two-part comprehensive search strategy based on previously published reviews (e.g., (1,14)) and our expertise to cover the full range of SSD/FSS while narrowing the search down to studies evaluating early psychological interventions (see Text in Section A, Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>, and the PROSPERO registration).

Moreover, L.B. and a research assistant conducted a backward search by independently checking the reference lists of eligible reports for further potentially eligible reports not identified during electronic database search.

Study Selection

Records were managed using EndNote (Version X9.2). Titles and abstracts were screened by L.B.

A duplicate screening of 30 randomly selected full-text records by LB and the research assistant resulted in an interrater agreement of 80% (Cohen's κ (30): $\kappa = 0.3$; prevalence-adjusted and bias-adjusted κ (31): $\kappa = 0.6$). Because we considered this interrater agreement to be sufficient, full-text screening was conducted by L.B. only, whereas ambiguous decisions regarding study selection were discussed with M.C.S.-M. The final study sample was checked for appropriateness by M.C.S.-M.

Data Collection Process

After the study selection process, we looked for corrections and errata of included reports to ensure data integrity. Duplicate reports were identified by L.B. and M.C.S.-M. and were treated as one study (32).

Data collection was conducted by L.B. using a standardized electronic form reviewed by L.L. and M.C.S.-M. priorly. L.B. discussed difficult coding decisions with M.C.S.-M. A subset of 10 studies was coded by L.B. and M.C.S.-M. in duplicate to establish interrater agreement in outcome data. A two-way random-effects single-rater intraclass correlation (33) indicated moderate to good reliability (ICC = 0.79, 95% confidence interval [CI] = 0.73–0.84) (34). Disagreements were mainly attributable to the selection of the control condition and post-treatment measurement. Disagreements were discussed to reach a consensus.

For 1 of 18 studies with insufficient information to compute effect sizes, we obtained all necessary data from the study authors.

Data Items

We extracted data on general study information, participants, intervention, outcomes, and further statistics relevant for cluster-randomized trials. Outcome data were extracted for baseline, post-treatment, and the longest follow-up measurement.

Length of follow-up was computed as the time frame between end of treatment and follow-up measurement. If length of follow-up varied between individuals within a study, we coded the shortest possible length of follow-up.

Risk of Bias in Individual Studies

Risk of bias was assessed at outcome level by L.B. We used the revised Cochrane risk-of-bias tool for individually randomized trials (RoB 2) (35). For cluster-randomized trials, we used the predecessor of this instrument (36) because it additionally addresses bias specific to this trial design. Risk-of-bias figures were created using the *robvis* package (37).

Summary Measures and Planned Methods of Analysis

Meta-analyses were performed, if at least three effect sizes were available for the respective analysis. Otherwise, we synthesized the studies narratively.

For each outcome at each time point (post-treatment versus follow-up), we computed random-effects meta-analyses within R (Version 3.6.0) (38) using the *metafor* package (39). Weights were computed using the inverse-variance method. Between-study variance (τ^2) was estimated using the method of restricted maximum likelihood. The 95% CIs were calculated via the Knapp-Hartung method (40,41). We also report 95% prediction intervals (42) describing the range in which effect sizes of future similar studies will probably fall.

For diagnostic status data, we report risk ratios (RRs), with numbers less than 1 representing more desirable results in the intervention group. For all other outcomes, we report Hedges' g (43), with positive numbers representing more desirable effects in the intervention group.

For cluster-randomized trials, we approximated correct effect size CIs by inflating standard errors (44). Analogous to van Dessel et al. (17), we imputed an intraclass correlation of 0.031 based on an estimation from Campbell et al. (45), when information on intraclass correlation was missing.

Risk of Bias Across Studies

We explored the range of possible outcomes when correcting for meta-bias by implementing the conditional precision-effect test and precision-effect estimate with standard errors (PET-PEESE) procedure (46) and a three-parameter selection model (3PSM) (47,48). In the conditional PET-PEESE procedure, effect sizes are regressed on their standard errors/variances (49). The resulting intercept serves as corrected effect estimate (46,49). As in Carter et al. (47), we implemented PET-PEESE with multiplicative error terms.

The 3PSM provides a corrected effect size estimate by specifying a selection model estimating the probability of observing effect sizes within prespecified p value ranges (48,50,51). In our analyses, the selection model was a step function consisting of two ranges with a one-sided p value of .025 as a cut point corresponding to the conventional significance criterion of $\alpha = .05$ (two-sided). We implemented the 3PSM approach by using the *weightR* package (52).

Additional Analyses

We conducted meta-regressions to analyze the impact of intervention intensity, mean duration of symptoms, type of participant population, and type of control group on the treatment effect. For effects at follow-up, we additionally investigated length of follow-up as a moderator.

To detect potential confounding between moderators, we examined their relationships descriptively. The relationship between intervention intensity and ordinal variables (type of participant population, type of control group) as well as the relationship among ordinal variables were examined using Cramer's V ($0 \hat{=}$ no association; $1 \hat{=}$ maximum dependence) (53,54).

The relationship between ordinal and metric variables (duration of symptoms, length of follow-up) was examined using Spearman's ρ . The relationship between intervention intensity and metric variables was assessed using biserial correlation (r_b), whereas the relationship between metric variables was quantified via Pearson's r .

Sensitivity Analyses

We investigated the robustness of the results to the method used for estimating between-study variance by repeating analyses, with τ^2 estimated via the two-step DerSimonian-Laird method (55,56). In addition, we repeated the analyses with individually randomized controlled trials, exclusively.

RESULTS

For more detailed information on outcome data and risk-of-bias ratings, see Tables and Figures in Section C, Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>.

Study Selection

The electronic database search identified 5499 different records (see Table in Section D, Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>). After full-text screening of 408 records and a backward search based on initial 29 eligible records, 37 records reporting 30 different studies were included (Table 1; see also Figure 1, for a flow diagram).

Study Characteristics

Of all included studies, 28 were individually randomized controlled trials, whereas two were cluster-randomized trials. Almost all studies were conducted in Europe, North America, or Australia.

In total, 11,342 individuals participated in the included studies (median = 176, range = 28–3851). Of these, 52.5% were female (2 studies with missing information). Twelve studies were conducted in a prevention population, whereas 18 studies were conducted in an early intervention population. The mean age ranged from 30.5 to 66 years (median = 41.3 years). Syndromes of interest were mostly pain related ($k = 25$), with chronic low back pain being investigated

TABLE 1. Characteristics of Included Studies

Study ^a	Design	n	Population ^b	Syndrome of Interest	Mean Duration of Symptoms	Interventions	No. Treatment Sessions
Bérubé et al., 2019	RCT	56	Prevention	Chronic postoperative pain	—	SC/TAU versus TAU + self-management intervention	7
Birch et al., 2020	RCT	67	Prevention	Chronic postoperative pain	—	SC/TAU versus TAU + CBT-oriented patient education	6–7
Bjørnnes et al., 2017	RCT	416	Prevention	Chronic postoperative pain	—	SC/TAU versus TAU + pain management education (educational booklet and discussion of its content during supportive follow-up phone call 10 days after surgery)	1
Cai et al., 2018	RCT	111	Prevention	Chronic postoperative pain	—	SC/TAU versus TAU + CBT	4
Dahl & Nilsson, 2001	RCT	29	Early intervention	Chronic pain	NA	SC/TAU versus CBT	11
Damush et al., 2003a (Damush et al., 2003b)	RCT	211	Early intervention	Chronic low back pain	NA	SC/TAU versus TAU + group-based self-management program	16
Ferrari et al., 2005	RCT	112	Prevention	Chronic whiplash syndrome	NA	SC/TAU versus TAU + whiplash prevention pamphlet	—
Gatchel et al., 2003	RCT	70	Early intervention	Chronic low back pain	NA	No treatment versus functional restoration program consisting of CBT, biofeedback, physical therapy, occupational therapy, and interdisciplinary team meetings	Up to 45
Gatchel et al., 2006 (Stowell et al., 2007)	RCT	101	Early intervention	Temporomandibular joint disorder	97.4 days	No treatment versus CBT and biofeedback	6
Gil-Jardiné et al., 2018	RCT	130	Prevention	Postconcussion syndrome	NA	SC/TAU versus TAU + reassurance (n = 38) or TAU + EMDR R-TEP (n = 34)	1
Hazard et al., 2000	RCT	489	Prevention	Chronic low back pain	NA	No treatment versus educational pamphlet	—
Irvine et al., 2015	RCT	597	Early intervention	Chronic low back pain	NA	No treatment versus online cognitive-behavioral tailored self-help program with weekly reminders for using the program	—

Author(s)	Study Design	Sample Size (n)	Intervention Type	Condition	Intervention (range)	Control (range)	Intervention Description	Wait-list Description	Number of Studies
Janse et al., 2016	RCT	100	Early intervention	Idiopathic chronic fatigue	Med = 5.5 yrs (range: 1-61 yrs)	Control: Med = 4 yrs (range: 0.5-26 yrs)	Wait-list versus tailored CBT-based guided self-instruction	—	
Karjalainen et al., 2004 (Karjalainen et al., 2003)	RCT	170	Early intervention	Chronic low back pain	NA	NA	SC/TAU versus TAU + advice and physiotherapeutic exercises (+ work site visit for n = 51)	1-2	
Kongsted et al., 2008	RCT	182	Prevention	Chronic whiplash syndrome	NA	NA	SC/TAU versus oral advice at home visit and list with key information	1	
Lamb et al., 2012 (Lamb et al., 2013)	CRT	3851	Prevention	Chronic whiplash syndrome	NA	NA	SC/TAU versus active management advice during consultation and educational pamphlet	1	
Linton & Andersson, 2000 (Linton & Nordin, 2006)	RCT	272	Early intervention	Chronic low back pain	NA	NA	SC/TAU versus TAU + group CBT	6	
Linton & Ryberg, 2001	RCT	253	Early intervention	Chronic low back pain, chronic neck pain	NA	NA	SC/TAU versus TAU + group CBT	6	
Mitchell & Carmen, 1994	RCT	542	Early intervention	Chronic pain	NA	NA	SC/TAU versus functional restoration program consisting of active exercise, functional simulation, CBT, psychoeducation, relaxation therapy, biofeedback and counseling; both individual- and group-format	Up to 280	
Newcomer et al., 2008	RCT	220	Early intervention	Chronic low back pain	NA	NA	SC/TAU versus educational videotape targeting beliefs and self-management	4	
Nyenhuis et al., 2013 (Nyenhuis et al., 2013)	RCT	304	Early intervention	Tinnitus	96 d ^c	96 d ^c	No treatment versus CBT-based Internet self-help or CBT-oriented bibliotherapy or group CBT	— (Internet self-help and bibliotherapy), 4 (group CBT)	
Riddle et al., 2019	RCT	402	Prevention	Chronic postoperative pain	NA	NA	Placebo versus TAU + CBT-based pain coping skills training	8	
Sanders et al., 2013	RCT	271	Early intervention	Temporomandibular joint disorder	NA	NA	Placebo versus CBT and biofeedback	6	
Sharpe et al., 2012 (study 1)	RCT	54	Prevention	Chronic pain	NA	NA	Placebo versus TAU + attention bias modification	1	

Continued on next page

TABLE 1. (Continued)

Study ^a	Design	n	Population ^b	Syndrome of Interest	Mean Duration of Symptoms	Interventions	No. Treatment Sessions
Silverberg et al., 2013	RCT	28	Prevention	Postconcussion syndrome	NA	SC/TAU versus TAU + CBT	6
Slater et al., 2009	RCT	67	Early intervention	Chronic low back pain	NA	Placebo versus TAU + behavioral pain management and rehabilitation intervention	6
Sterling et al., 2019	RCT	108	Early intervention	Chronic whiplash syndrome	NA	SC/TAU versus TAU + exercise and CBT-oriented stress inoculation training	6
Toft et al., 2010	CRT	111	Early intervention	Any somatoform disorder (ICD-10: F44.4–F48.0)	NA	SC/TAU versus PCP training in assessment and treatment of FSS	7
Traeger et al., 2019	RCT	202	Early intervention	Chronic low back pain	13 days	Placebo versus TAU + intensive patient education	2
Whitfill et al., 2010 (Rogerson et al., 2010)	RCT	142	Early intervention	Chronic low back pain	NA	SC/TAU versus functional restoration program consisting of CBT, biofeedback, physical therapy, occupational therapy and interdisciplinary team meetings	Up to 24

RCT = randomized controlled trial; — (dash) = not applicable to the respective study; SC/TAU = standard care/treatment as usual; CBT = cognitive-behavioral therapy; NA = missing data; EMDR = eye movement desensitization and reprocessing; R-TEP = recent traumatic episode protocol; Med = Median; CRT = cluster-randomized controlled trial; PCP = primary care physician; FSS = functional somatic syndromes.

^aReferences in parentheses indicate duplicate reports. For full references, see Text in Section E, Supplemental Digital Content (<http://links.lww.com/PSYMED/A785>).

^bPopulation coded in accordance with our participant eligibility criteria (see Eligibility Criteria, for an explanation).

^cIn this study, mean symptom duration was reported in months. Mean symptom duration in days was calculated by assuming a month has 30 days.

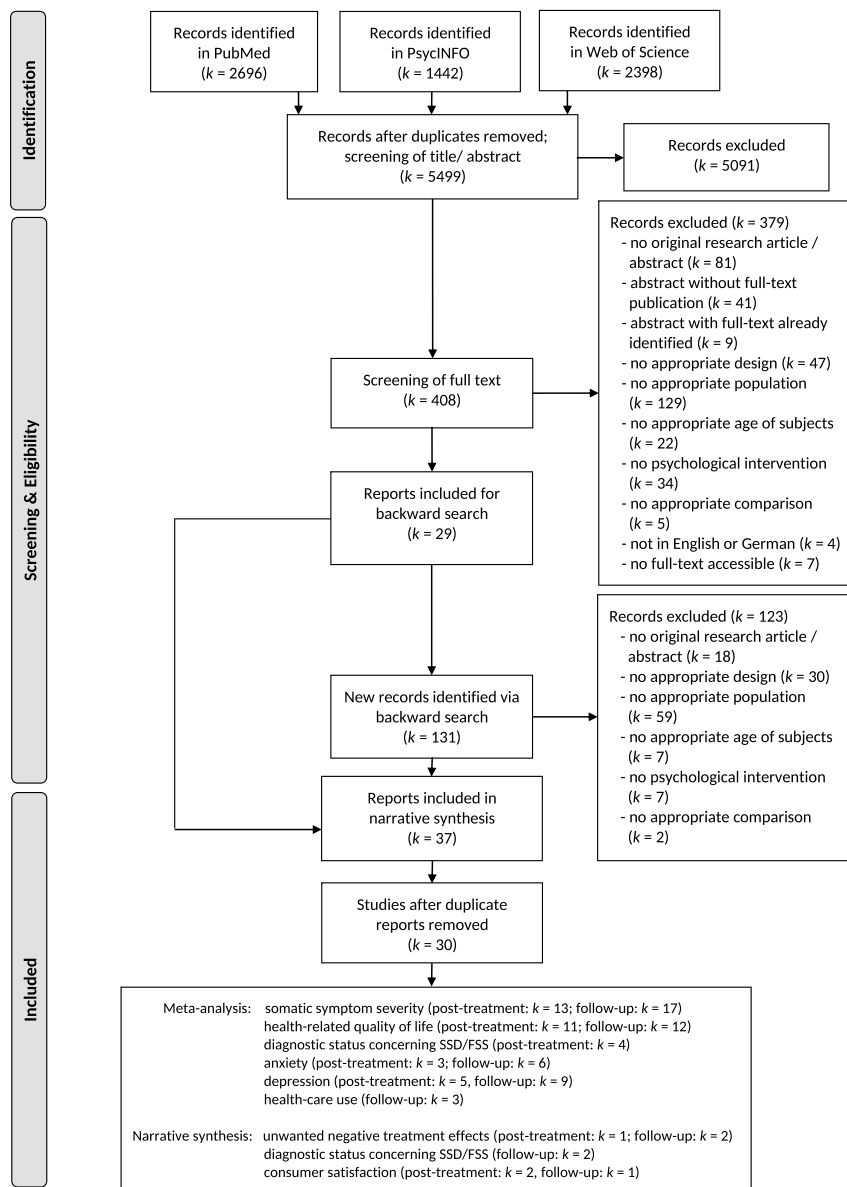


FIGURE 1. PRISMA flow diagram (57). Numbers of excluded records also include records for which the respective information was unclear. Numbers of included studies in the narrative and meta-analytic syntheses represent the numbers of studies with available data. Seven full texts could not be obtained despite contacting authors.

most frequently ($k = 11$). Mean symptom duration at baseline was reported in four studies and ranged from 13 days to 5.5 years.

Interventions in the treatment group were diverse. One study evaluated a training for general practitioners, whereas 29 studies evaluated patient-directed interventions. Of these, 24 studies consisted of interventions comprising contact with a health care professional. In the remaining studies, interventions were completely based on written materials ($k = 2$), in self-help format ($k = 1$), video-based ($k = 1$), or included a computer-based paradigm (attention bias modification, $k = 1$). Five studies evaluated early psychological interventions in combination with other treatments, for example, physiotherapeutic exercises.

Eight of the 29 studies with patient-directed interventions had an educational focus (i.e., predominantly providing information),

whereas 21 studies had a therapeutic focus (i.e., predominantly targeting change in psychological processes). In the latter group, interventions in 16 studies included a cognitive-behavioral therapy-based component, whereas the other 5 studies evaluated either eye movement desensitization and reprocessing (EMDR) as well as reassurance, self-management ($k = 2$), stress inoculation training, or attention bias modification. Cognitive-behavioral therapy as such was implemented in 14 studies, including an in-group format ($k = 3$), in combination with biofeedback ($k = 2$), or as part of multidisciplinary treatment programs ($k = 3$).

In four studies, the number of treatment sessions could not be determined because of the nature of interventions (e.g., self-help format, written material). In the remaining studies, the median number of treatment sessions was 6, ranging from 1 to 280 (the

latter being an outlier of one study of multidisciplinary inpatient treatment of 8–12 weeks in duration (58)).

Early psychological interventions were compared with SC/TAU ($k = 19$), no treatment ($k = 5$), psychological placebo treatment ($k = 5$), and wait-list control ($k = 1$).

Primary Outcomes

Somatic Symptom Severity

Post-treatment

Of 19 studies measuring somatic symptom severity post-treatment, effect size data were available for 13 studies ($n = 2031$). There was a nonsignificant effect ($g = 0.11 [-0.079 \text{ to } 0.3]$; Figure 2). Heterogeneity was significantly different from zero ($Q(12) = 38.3, p < .001$), and inconsistency was moderate to considerable ($I^2 = 67.7\% [33.3\% \text{ to } 88.8\%]$). The resulting 95% prediction interval ranged from -0.46 to 0.68 . Risk of bias is depicted in Figure 2.

The PET-PEESE revealed a corrected effect estimate of $g = 0.008$ ($-0.51 \text{ to } 0.52$). The 3PSM revealed a corrected effect estimate of $g = -0.023$ ($-0.2 \text{ to } 0.15$). A likelihood ratio test did not reveal a significantly better fit of the 3PSM to the data ($\chi^2(1) = 2.68, p = .10$).

Follow-Up

Of 24 studies measuring somatic symptom severity at follow-up, effect size data were available for 17 studies ($n = 2438$). Follow-up length ranged from 1.5 to 24 months (median = 9.5). There was a small, significant positive effect ($g = 0.25 [0.096 \text{ to } 0.41]$; Figure 3). Heterogeneity was significantly different from zero ($Q(16) = 37.4, p = .002$), and inconsistency was small to considerable ($I^2 = 58.4\% [22.9\% \text{ to } 85.9\%]$). The resulting 95% prediction interval ranged from -0.22 to 0.72 . Risk of bias is depicted in Figure 3.

The PET-PEESE revealed a corrected effect estimate of $g = 0.04 [-0.32 \text{ to } 0.4]$. The 3PSM revealed a corrected effect estimate of $g = 0.14 [-0.041 \text{ to } 0.31]$. A likelihood ratio test did not reveal a significantly better fit of the 3PSM to the data ($\chi^2(1) = 2.01, p = .16$).

Health-Related Quality of Life

Post-treatment

Of 17 studies measuring HrQoL post-treatment, effect size data were available for 11 studies ($n = 4498$). There was a nonsignificant effect ($g = 0.13 [-0.07 \text{ to } 0.33]$; Figure 4). Heterogeneity

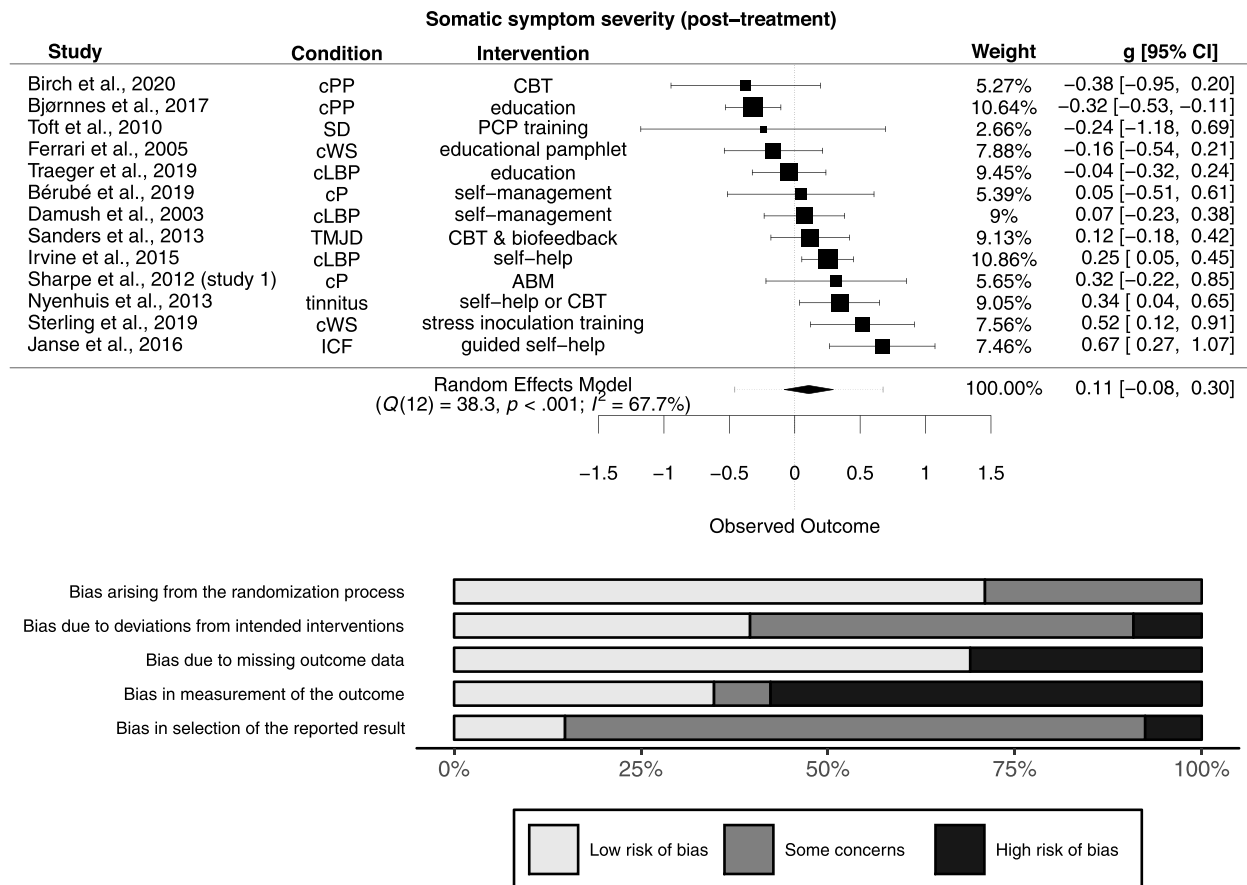


FIGURE 2. Forest plot and risk of bias inherent in the summary effect for somatic symptom severity (post-treatment). $g > 0$ indicates more favorable outcomes in the intervention group. ABM = attention bias modification; CBT = cognitive-behavioral therapy; cLBP = chronic low back pain; cP = chronic pain; cPP = chronic postoperative pain; cWS = chronic whiplash syndrome; ICF = idiopathic chronic fatigue; PCP = primary care physician; SD = somatoform disorder; TMJD = temporomandibular joint disorder. Study-level biases are weighted according to the meta-analytic weights. One cluster-randomized study was included in this meta-analysis (59). There was a high risk of bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomization in this study (not depicted).

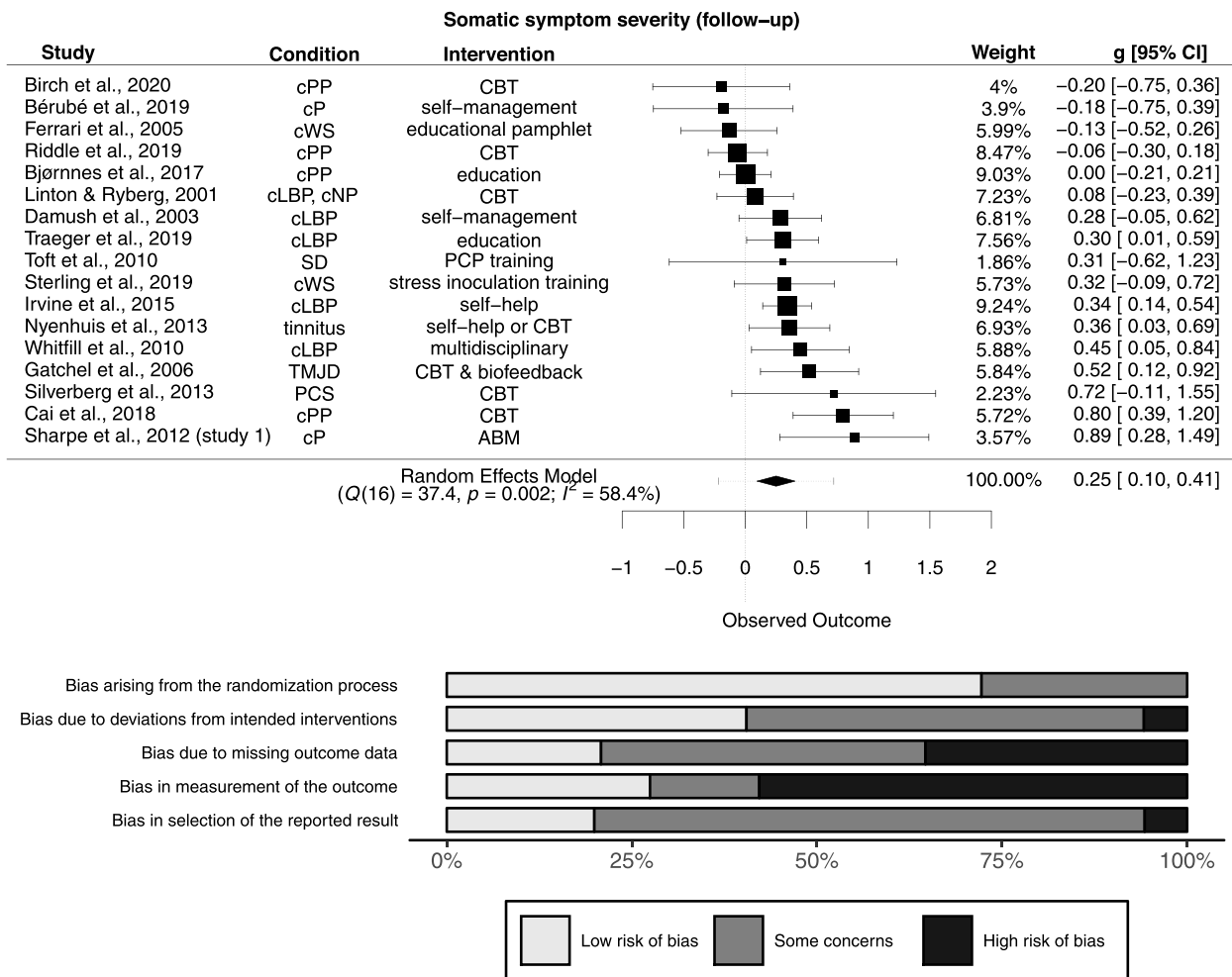


FIGURE 3. Forest plot and risk of bias inherent in the summary effect for somatic symptom severity (follow-up). $g > 0$ indicates more favorable outcomes in the intervention group. ABM = attention bias modification; CBT = cognitive-behavioral therapy; cLBP = chronic low back pain; cNP = chronic neck pain; cP = chronic pain; cPP = chronic postoperative pain; cWS = chronic whiplash syndrome; PCP = primary care physician; PCS = postconcussion syndrome; SD = somatoform disorder; TMJD = temporomandibular joint disorder. Study-level biases are weighted according to the meta-analytic weights. One cluster-randomized study was included in this meta-analysis (59). There was a high risk of bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomization in this study (not depicted).

was significantly different from zero ($Q(10) = 19.9, p = .03$), and inconsistency was small to considerable ($I^2 = 51.4\%$ [0% to 88.8%]). The resulting 95% prediction interval ranged from -0.34 to 0.60. Risk of bias is depicted in Figure 4.

The PET-PEESE revealed a corrected effect estimate of $g = 0.23$ [-0.013 to 0.48]. The 3PSM revealed a corrected effect estimate of $g = 0.055$ [-0.13 to 0.24]. A likelihood ratio test did not reveal a significantly better fit of the 3PSM to the data ($\chi^2(1) = 1.79, p = .18$).

Follow-Up

Of 18 studies measuring HrQoL at follow-up, effect size data were available for 12 studies ($n = 1681$). Follow-up length ranged from 2 to 24 months (median = 10). There was a nonsignificant effect ($g = 0.12$ [-0.001 to 0.25]; Figure 5). Heterogeneity was not significantly different from zero ($Q(11) = 13.2, p = .28$), and inconsistency was small to substantial ($I^2 = 23.1\%$ [0% to 72.7%]). The resulting 95% prediction interval ranged from -0.13 to 0.37. Risk of bias is depicted in Figure 5.

The PET-PEESE revealed a corrected effect estimate of $g = 0.18$ [0.002 to 0.36]. The 3PSM revealed a corrected effect estimate of $g = 0.16$ [-0.004 to 0.32]. A likelihood ratio test did not reveal a significantly better fit of the 3PSM to the data ($\chi^2(1) = 0.45, p = .5$).

Secondary Outcomes

We could detect significant effects of early psychological interventions on depression posttreatment ($g = 0.12$ [0.03 to 0.2], $k = 5$) and on health care utilization at follow-up ($g = 0.31$ [0.18 to 0.44], $k = 3$). While the effect on depression post-treatment turned nonsignificant when applying PET-PEESE ($g = 0.12$ [-0.4 to 0.64]) and was robust to the 3PSM ($g = 0.17$ [0.046 to 0.29]; $\chi^2(1) = 1.32, p = .25$), the effect of health care utilization at follow-up was robust to PET-PEESE ($g = 0.26$ [0.15 to 0.38]) and increased in the 3PSM ($g = 0.82, \chi^2(1) = 5.75, p = .016$, no CI available due to model convergence problems). For all other outcomes, there were no significant effects, or there were too

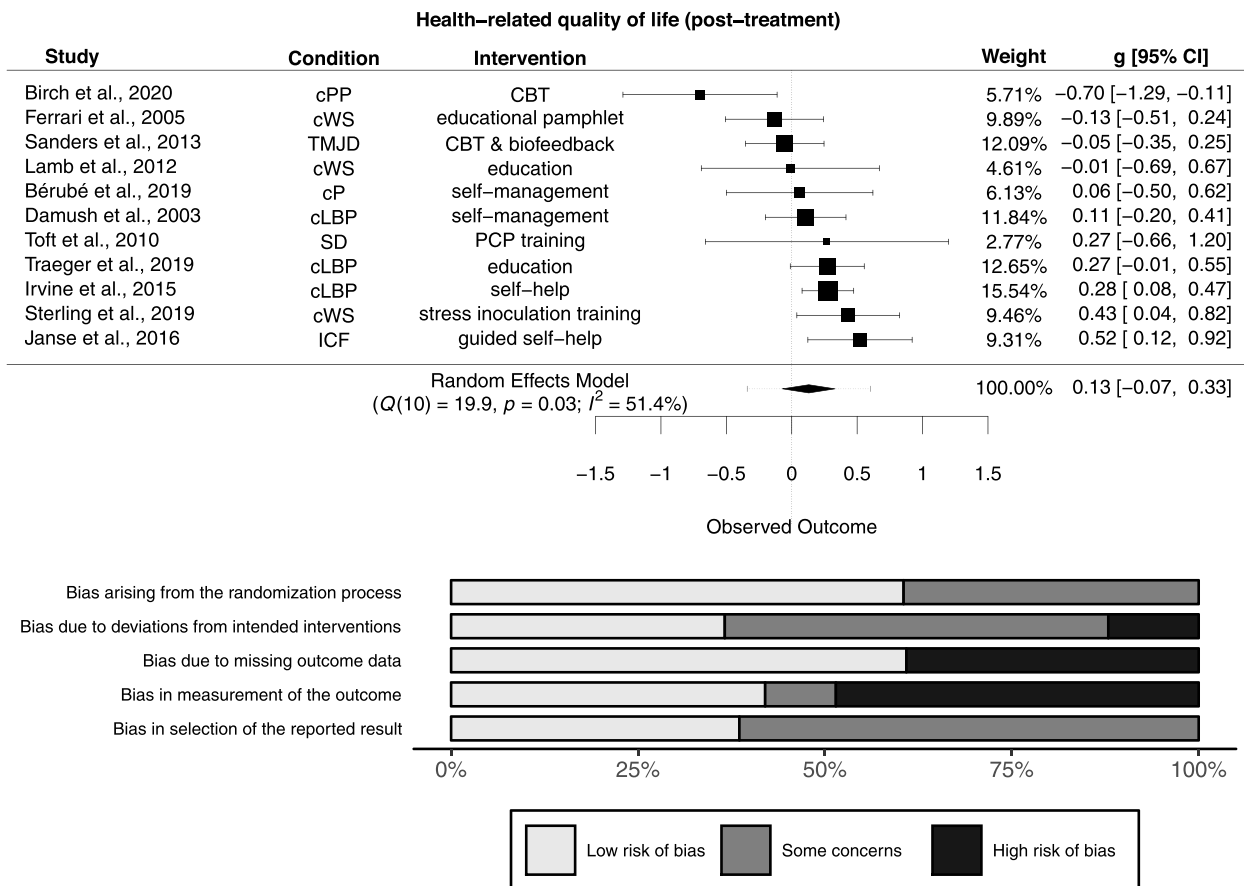


FIGURE 4. Forest plot and risk of bias inherent in the summary effect for health-related quality of life (post-treatment). $g > 0$ indicates more favorable outcomes in the intervention group. CBT = cognitive-behavioral therapy; cLBP = chronic low back pain; cP = chronic pain; cPP = chronic postoperative pain; cWS = chronic whiplash syndrome; ICF = idiopathic chronic fatigue; PCP = primary care physician; SD = somatoform disorder; TMJD = temporomandibular joint disorder. Study-level biases are weighted according to the meta-analytic weights. Two cluster-randomized studies were included in this meta-analysis (59,60). Whereas the study by Lamb et al. (60) was at low risk of bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomization, the study by Toft et al. (59) was at high risk (not depicted).

few data available to conduct meta-analyses. For full results concerning the secondary outcomes, see Text in Section F, Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>.

Additional Analyses

Type of population significantly moderated the effect in three outcomes: somatic symptom severity post-treatment ($F(1,11) = 7.14, p = .022, R^2 = 63.1%$), HrQoL post-treatment ($F(1,9) = 6.91, p = .027, R^2 = 80%$), and HrQoL at follow-up ($F(1,10) = 6.14, p = .033, R^2 = 100%$). Post-hoc analyses revealed nonsignificant effects in prevention populations (g 's = -0.18 to -0.028) and significant effects in early intervention populations (g 's = 0.21 to 0.24).

Descriptive analyses of these outcomes showed confounding of type of population with intervention intensity (V 's = 0.29–0.39), with high intensity interventions tending to be overrepresented in studies with early intervention populations. Furthermore, there was confounding with type of control group (V 's = 0.36–0.57), with prevention populations being investigated in studies with SC/TAU or placebo controls, only. For HrQoL at follow-up, the large positive rank correlation indicated longer follow-ups in early intervention populations ($\rho = 0.52$). See Text in Section G,

Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>, for full results of additional analyses.

Sensitivity Analyses

When rerunning analyses with individually randomized trials only, the pattern of results did not change. When repeating analyses with the two-step DerSimonian-Laird estimator for τ^2 , the effect on HrQoL at follow-up became significant ($g = 0.13 [0.007–0.25]$). The remaining pattern of results was unchanged.

In a post-hoc sensitivity analysis, we repeated our main analyses without the study by Janse et al. (61). Although individuals were eligible for this review by suffering from a subthreshold SSD/FFS, reported symptom durations of 4 to 5.5 years made the earliness of the intervention disputable. The reanalysis did not considerably change the summary effects. For full results of sensitivity analyses, see Text in Sections H, I, and J, Supplemental Digital Content, <http://links.lww.com/PSYMED/A804>.

DISCUSSION

We conducted a systematic review and meta-analysis to assess the efficacy of early psychological interventions in preventing and

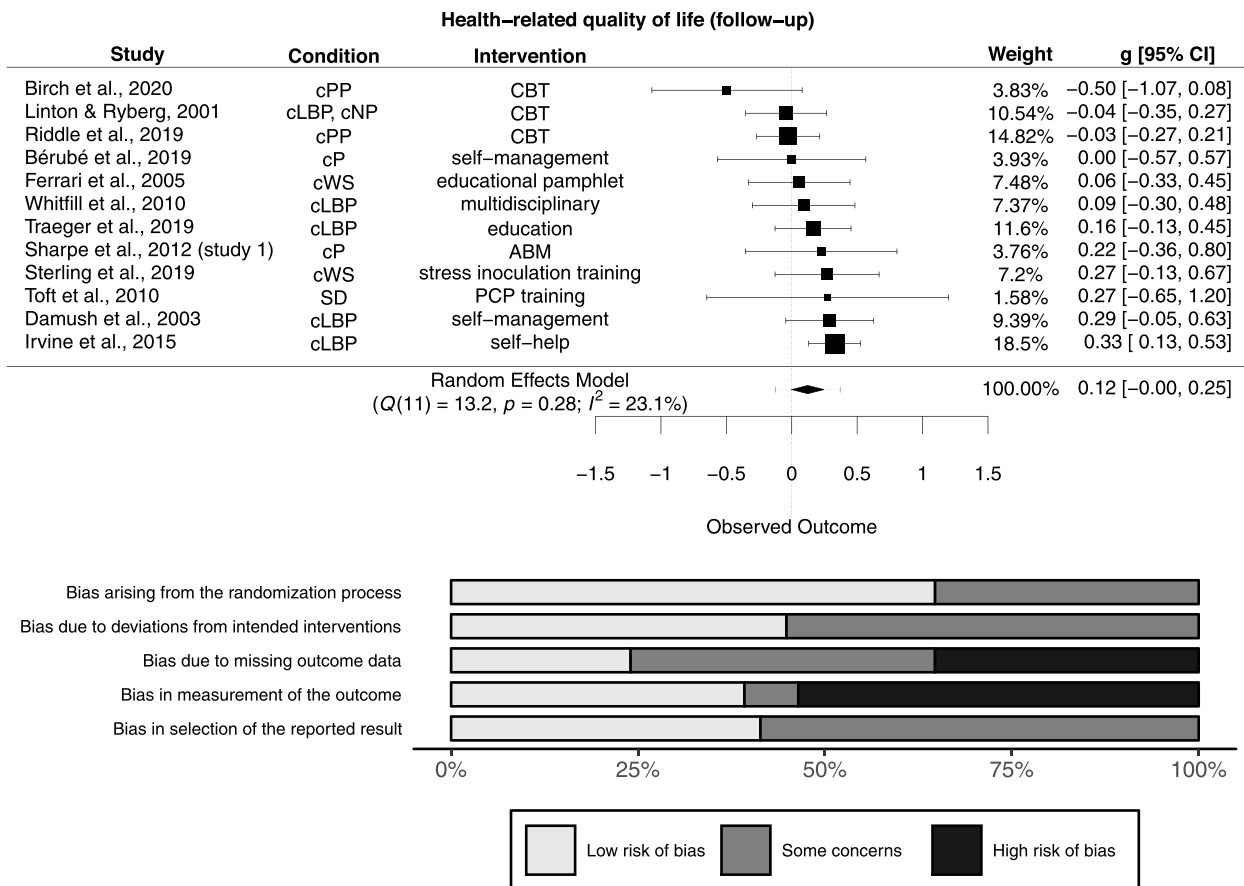


FIGURE 5. Forest plot and risk of bias inherent in the summary effect of health-related quality of life (follow-up). $g > 0$ indicates more favorable outcomes in the intervention group. ABM = attention bias modification; CBT = cognitive-behavioral therapy; cLBP = chronic low back pain; cNP = chronic neck pain; cP = chronic pain; cPP = chronic postoperative pain; cWS = chronic whiplash syndrome; ICF = idiopathic chronic fatigue; PCP = primary care physician; SD = somatoform disorder; TMJD = temporomandibular joint disorder. Study-level biases are weighted according to the meta-analytic weights. One cluster-randomized study was included in this meta-analysis (59). There was a high risk of bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomization in this study (not depicted).

treating SSD/FSS in adults compared with inactive controls, SC/TAU, or placebo. Assuming that the DUI plays a relevant role in determining treatment efficacy (e.g., (19)), we hypothesized that early psychological interventions would show positive effects. Overall, we were unable to find reliable evidence supporting our hypotheses.

In most of the cases, our meta-analyses either revealed no significant effects on core outcomes in SSD/FSS research (28) or could not be conducted, as too few data were available. However, for depression post-treatment as well as somatic symptom severity, health care utilization, and HrQoL (only sensitivity analysis) at follow-up, our meta-analyses revealed significant positive effects. These effects resemble the small- to medium-sized effects found in previous reviews (e.g., (16,17)). Importantly, there was a high risk of bias within studies, and the effects were sensitive to corrections for meta-bias. Consequently, the effects should not be considered as valid estimates of the population effect (48,62).

For somatic symptom severity and HrQoL post-treatment as well as for HrQoL at follow-up, we found a significant moderation by type of population indicating no effects in studies with prevention populations and small effects in studies with early intervention populations. On the one hand, this moderation might be driven by

higher rates of spontaneous remissions in studies with prevention populations compared with studies with early intervention populations, concealing intervention effects. On the other hand, this moderation might be attributable to more intensive interventions and less active comparators in studies with early interventions, as these variables were confounded with type of population. Of note, these findings are based on a large number of statistical tests and a low number of studies. Thus, the moderation should be interpreted with caution.

Despite the uncertainty inherent in our statistical results, our review provides a valuable insight into the shortcomings of previous research on early psychological interventions for SSD/FSS. In the vast majority of included studies, pain-related conditions, in particular low back pain, were examined. Relevantly, we could not detect any studies targeting cardiopulmonary-autonomic and gastrointestinal SSD/FSS subtypes (63). Furthermore, we did not detect any study meeting our first presentation criterion. This could either result from using inadequate search terms or rather reflect the lack of such studies, as identifying SSD/FSS in primary care is very challenging because of difficulties in physician-patient interaction or the fluctuating course of SSD/FSS symptoms (24–26). Although we aimed to capture a population very early in SSD/FSS

development, we were confronted with the lack of current consensus regarding the definition of “early interventions” in the field of SSD/FSS. As an approximation, we developed three different inclusion criteria to capture the population of interest as comprehensively as possible. However, this might have introduced considerable heterogeneity, thereby impeding straightforward conclusions based on this review.

Because 53% of individuals included in this review identified as female, our review might not mirror the typical sex distribution in SSD/FSS studies in which females often represent the majority (e.g., (17)). This difference in sex distribution might be attributable to the specific populations targeted in the included studies, for example, workers with an occupational injury, in which sex distribution was more balanced or male individuals even in the majority. Although these samples comply with our aim to include individuals at risk or at an early stage of SSD/FSS, this difference in sex distribution might limit the generalizability of our findings.

Interventions were diverse and ranged from educational pamphlets to multidisciplinary treatment programs consisting of up to 280 treatment sessions. This diversity is probably a result of the various approaches adopted by the different disciplines involved in the treatment of SSD/FSS. Beyond that, interventions were compared with placebo controls in a few studies only, impeding the identification of active treatment components.

Limitations

Although, in general, the included studies limited the maximum symptom duration in their eligibility criteria, only four studies explicitly reported mean symptom durations. Because SSD/FSS symptoms wax and wane (24,64) and changes in the SSD/FSS subtype occur (26), individuals might have fulfilled the study-specific symptom duration criteria despite actually having persistent symptoms. In fact, the study by Janse et al. (61) reported a median symptom duration of 4 to 5.5 years. We acknowledge that the fluctuating nature of SSD/FSS symptoms makes it difficult to draw a precise line between early and usual interventions. We tried to manage this challenge by using comprehensive inclusion criteria allowing an examination of different conceptualizations of early interventions. Nonetheless, it is possible that the primary studies included in this review did not adequately capture patients at an early stage of SSD/FSS development.

The interpretation of our findings may be hampered by the methodology we used. The interrater agreement in study selection and data extraction was far from optimal questioning the appropriateness as well as the completeness of our study sample. Double coding the full data set would have represented a more effective safeguard against errors in study selection and data extraction. Furthermore, the inclusion of studies evaluating psychological interventions in combination with other treatments such as physiotherapy might have biased results in a more positive direction.

Beyond that, our analyses are based on a low number of effects, even impeding meta-analytic integration in some outcomes. Missing outcome data further reduced the number of available effects by one quarter. This selective availability of outcome data may indicate a reporting bias withholding nonsignificant effects (65). These circumstances have probably biased the results in a more positive direction, hampered the validity of inconsistency estimates

and tests of heterogeneity (66), and led to false-negative results in the meta-bias corrections (62,67,68).

CONCLUSIONS

Our review suggests that the area of early psychological interventions in SSD/FSS is highly underresearched and not well elaborated on a conceptual level, which is mirrored in an unsystematic and heterogeneous research output. We perceive an urgent need in the following four major research efforts: First, innovative treatment approaches tailored to at-risk or newly affected populations should be developed, crossing the borders of medical specialties and diagnostic constructs. These approaches should direct attention toward common transsymptomatic mechanisms and processes of change (69) to target empirically derived psychosocial and biological factors as well as their interactions, thereby elucidating their role in mediating or moderating treatment response. Moreover, a better biopsychosocial mechanistic understanding of how acute events lead to persistent somatic symptoms can help to develop more targeted early psychological interventions in the future. Second, the symptom history of the included individuals should be thoroughly assessed aiming at capturing individuals with SSD/FSS at an earlier stage. Third, effects on core outcomes in SSD/FSS (28) should be fully disclosed, as suggested by the European Network to Improve Diagnosis, Treatment and Health Care for Patients with Persistent Somatic Symptoms (EURONET-SOMA) (70). Fourth, placebo-controlled designs should be used, allowing for an identification of specific active treatment components.

We would like to highlight two early intervention approaches that we perceive to be especially promising in addressing the lack of interventions targeting at-risk or newly affected populations as well as adopting a transsymptomatic approach. First, we think that interventions at primary care level (e.g., (59)) might effectively reach persons suffering from a wide range of SSD/FSS subtypes at an early stage (71), as many individuals with different symptoms seek help in primary care in the first place, already. Second, (guided) self-help interventions (e.g., (61)) might allow for resource-saving treatment and may be constructed in line with a transsymptomatic approach, thereby increasing the number of potential recipients. Their efficacy has been demonstrated in adults (18); however, they need to find ways of how to reach the affected individuals early, and their potential in young people further needs to be explored (72).

Based on the limitations of included primary studies and of our methodology, the validity of the effect sizes calculated in this review is questionable. Accordingly, the question whether the DUI is responsible for the low to medium effects attained by standard psychological treatments in SSD/FSS must remain unanswered. Based on research in other psychological disorders (e.g., (20)) and knowledge about disorder-maintaining processes in SSD/FSS (73), we still support the notion that a high DUI negatively affects treatment outcome in SSD/FSS. This notion is mirrored in current recommendations and guidelines promoting stepped, collaborative, and coordinated health care in the treatment of SSD/FSS (1,14). Stepped, collaborative health networks such as the Sofu-Net may increase access to psychological treatments and accordingly have the potential to reduce the DUI (e.g., (27,71)). Therefore, practitioners should continue to aim at treating SSD/FSS as early as possible.

To the best of our knowledge, our preregistered review was the first to systematically examine early psychological interventions

for such a broad range of SSD/FSS subtypes. As the efficacy of early psychological interventions and the role of the DUI in determining outcome remain unclear, more high-quality primary studies are needed. We hope that our findings inform future research elucidating the different factors involved in the development, maintenance, and treatment response of SSD/FSS to reduce suffering from this highly prevalent, burdensome condition.

This study was inspired and influenced by the scientific exchange within the European Network to Improve Diagnosis, Treatment and Health Care for Patients with Persistent Somatic Symptoms (EURONET-SOMA; <https://www.euronet-soma.eu/>). We would like to thank Maja Glahn for her very competent and reliable help in screening records and evaluating interrater agreement. We are grateful to all researchers providing us with additional information and data to their studies included in this review.

Source of Funding and Conflicts of Interest: The authors have no conflicts of interest and no sources of funding to declare.

REFERENCES

- Henningsen P, Zipfel S, Sattel H, Creed F. Management of functional somatic syndromes and bodily distress. *Psychother Psychosom* 2018;87:12–31.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Pub; 2013.
- Fink P. Syndromes of bodily distress or functional somatic syndromes—where are we heading. Lecture on the occasion of receiving the Alison Creed award 2017. *J Psychosom Res* 2017;97:127–30.
- Burton C, Fink P, Henningsen P, Löwe B, Rief W, EURONET-SOMA Group. Functional somatic disorders: discussion paper for a new common classification for research and clinical use. *BMC Med* 2020;18:34.
- Fink P, Schröder A. One single diagnosis, bodily distress syndrome, succeeded to capture 10 diagnostic categories of functional somatic syndromes and somatoform disorders. *J Psychosom Res* 2010;68:415–26.
- Nimnuan C, Rabe-Hesketh S, Wessely S, Hotopf M. How many functional somatic syndromes? *J Psychosom Res* 2001;51:549–57.
- Petersen MW, Schröder A, Jørgensen T, Ørnbøl E, Dantoft TM, Eliassen M, Carstensen TW, Falgaard Epløv L, Fink P. Prevalence of functional somatic syndromes and bodily distress syndrome in the Danish population: the DanFunD study. *Scand J Public Health* 2019;1–10.
- Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, Olesen J, Allgulander C, Alonso J, Faravelli C, Fratiglioni L, Jennum P, Lieb R, Maercker A, van Os J, Preisig M, Salvador-Carulla L, Simon R, Steinhausen H-C. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 2011;21:655–79.
- Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. *Dtsch Arztebl Int* 2015;112:279–87.
- Henningsen P, Zimmermann T, Sattel H. Medically unexplained physical symptoms, anxiety, and depression: a meta-analytic review. *Psychosom Med* 2003;65:528–33.
- Joustra ML, Janssens KA, Bültmann U, Rosmalen JG. Functional limitations in functional somatic syndromes and well-defined medical diseases. Results from the general population cohort LifeLines. *J Psychosom Res* 2015;79:94–9.
- Andersen NL, Epløv L, Andersen JT, Hjorthøj CR, Birket-Smith M. Health care use by patients with somatoform disorders: a register-based follow-up study. *Psychosomatics* 2013;54:132–41.
- Fink P. Surgery and medical treatment in persistent somatizing patients. *J Psychosom Res* 1992;36:439–47.
- Roenneberg C, Hausteiner-Wiehle C, Schäfer R, Sattel H, Henningsen P. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF): S3 Leitlinie “Funktionelle Körperbeschwerden”. 2018. Available at: https://www.awmf.org/uploads/tx_szleitlinien/051-0011_S3_Funktionelle_Koerperbeschwerden_2018-11.pdf. Accessed September 25, 2019.
- Abbass A, Town J, Holmes H, Luyten P, Cooper A, Russell L, Lumley MA, Schubiner H, Allinson J, Bernier D, De Meulemeester C, Kroenke K, Kisely S. Short-term psychodynamic psychotherapy for functional somatic disorders: a meta-analysis of randomized controlled trials. *Psychother Psychosom* 2020;89:363–70.
- Kleinstäuber M, Wittthöft M, Hiller W. Efficacy of short-term psychotherapy for multiple medically unexplained physical symptoms: a meta-analysis. *Clin Psychol Rev* 2011;31:146–60.
- van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, Numans ME, van der Horst HE, van Marwijk H. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. *Cochrane Database Syst Rev* 2014;CD011142.
- van Gils A, Schoevers RA, Bonvanie IJ, Gelauff JM, Roest AM, Rosmalen JG. Self-help for medically unexplained symptoms: a systematic review and meta-analysis. *Psychosom Med* 2016;78:728–39.
- Herzog A, Shedden-Mora MC, Jordan P, Löwe B. Duration of untreated illness in patients with somatoform disorders. *J Psychosom Res* 2018;107:1–6.
- Dell’Osso B, Glick ID, Baldwin DS, Altamura AC. Can long-term outcomes be improved by shortening the duration of untreated illness in psychiatric disorders? A Conceptual framework. *Psychopathology* 2013;46:14–21.
- Koelen JA, Houtveen JH, Abbass A, Luyten P, Eurelings-Bontekoe EH, Van Broeckhuysen-Kloth SA, Bühring ME, Geenen R. Effectiveness of psychotherapy for severe somatoform disorder: meta-analysis. *Br J Psychiatry* 2014;204:12–9.
- Löwe B, Lohse A, Andresen V, Vettorazzi E, Rose M, Broicher W. The development of irritable bowel syndrome: a prospective community-based cohort study. *Am J Gastroenterol* 2016;111:1320–9.
- Rask CU, Olsen EM, Elberling H, Christensen MF, Ørnbøl E, Fink P, Thomsen PH, Skovgaard AM. Functional somatic symptoms and associated impairment in 5–7-year-old children: the Copenhagen Child Cohort 2000. *Eur J Epidemiol* 2009;24:625–34.
- Claassen-van Dessel N, van der Wouden JC, Hoekstra T, Dekker J, van der Horst HE. The 2-year course of medically unexplained physical symptoms (MUPS) in terms of symptom severity and functional status: results of the PROSPECTS cohort study. *J Psychosom Res* 2018;104:76–87.
- Leiknes KA, Finset A, Moum T, Sandanger I. Overlap, comorbidity, and stability of somatoform disorders and the use of current versus lifetime criteria. *Psychosomatics* 2008;49:152–62.
- Murray AM, Toussaint A, Althaus A, Löwe B. The challenge of diagnosing non-specific, functional, and somatoform disorders: a systematic review of barriers to diagnosis in primary care. *J Psychosom Res* 2016;80:1–10.
- Shedden-Mora MC, Groß B, Lau K, Gumz A, Wegscheider K, Löwe B. Collaborative stepped care for somatoform disorders: a pre-post-intervention study in primary care. *J Psychosom Res* 2016;80:23–30.
- Rief W, Burton C, Frosthalm L, Henningsen P, Kleinstäuber M, Kop WJ, Löwe B, Martin A, Malt U, Rosmalen J, Schröder A, Shedden-Mora M, Toussaint A, van der Feltz-Cornelis C, EURONET-SOMA Group. Core outcome domains for clinical trials on somatic symptom disorder, bodily distress disorder, and functional somatic syndromes: european network on somatic symptom disorders recommendations. *Psychosom Med* 2017;79:1008–15.
- Barnsley L, Lord S, Bogduk N. Whiplash injury. *Pain* 1994;58:283–307.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
- Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;46:423–9.
- von Elm E, Poglía G, Walder B, Tramer MR. Different patterns of duplicate publication: an analysis of articles used in systematic reviews. *JAMA* 2004;291:974–80.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
- Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Juni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
- Eldridge SM, Campbell MK, Dahota A, Giraudeau B, Higgins JPT, Reeves B, Siegfried N. Revised Cochrane risk of bias tool for randomized trials (RoB 2.0). Additional considerations for cluster-randomized trials. 2016. Available at: <https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/archive-rob-2-0-cluster-randomized-trials-2016>. Accessed March 30, 2020.
- McGuinness LA. robvis: An R package and web application for visualising risk-of-bias assessments (Version 0.3.0). 2019.
- R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2019.
- Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.
- IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014;14:25.
- Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med* 2003;22:2693–710.
- IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016;6:e010247.
- Hedges LV. Distribution theory for Glass’s estimator of effect size and related estimators. *J Educ Behav Stat* 1981;6:107–28.
- Higgins JP, Deeks JJ, Altman DG. Special topics in statistics. In: Higgins JP, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [Updated March 2011]. London, United Kingdom: The Cochrane Collaboration; 2011:481–529.
- Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intraclass correlation coefficient in cluster randomized trials: the case of implementation research. *Clin Trials* 2005;2:99–107.

46. Stanley TD, Doucouliagos H. Meta-regression approximations to reduce publication selection bias. *Res Synth Methods* 2014;5:60–78.
47. Carter EC, Schönbrodt FD, Gervais WM, Hilgard J. Correcting for bias in psychology: a comparison of meta-analytic methods. *Adv Methods Pract Psychol Sci* 2019;2:115–44.
48. McShane BB, Böckenholt U, Hansen KT. Adjusting for publication bias in meta-analysis: an evaluation of selection methods and some cautionary notes. *Perspect Psychol Sci* 2016;11:730–49.
49. Stanley TD. Limitations of PET-PEESE and other meta-analysis methods. *Soc Psychol Pers Sci* 2017;8:581–91.
50. Iyengar S, Greenhouse JB. Selection models and the file drawer problem. *Stat Sci* 1988;3:109–17.
51. Vevea JL, Hedges LV. A general linear model for estimating effect size in the presence of publication bias. *Psychometrika* 1995;60:419–35.
52. Coburn KM, Vevea JL. weightr: estimating weight-function models for publication bias (Version 2.0.1). 2019.
53. Acock AC, Stavig GR. A measure of association for nonparametric statistics. *Soc Forces* 1979;57:1381–6.
54. Sakoda JM. Measures of association for multivariate contingency tables. *Proc Soc Stat Sect (Part III)* 1977;777–80.
55. DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. *Contemp Clin Trials* 2007;28:105–14.
56. Viechtbauer W. DerSimonian and Kacker (2007). The metafor Package. A Meta-Analysis Package for R. 2014. Available at: <http://www.metafor-project.org/doku.php/analyses:dersimonian2007>. Accessed May 25, 2020.
57. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
58. Mitchell RI, Carmen GM. The functional restoration approach to the treatment of chronic pain in patients with soft tissue and back injuries. *Spine (Phila Pa 1976)* 1994;19:633–42.
59. Toft T, Rosendal M, Ømbøl E, Olesen F, Frostholm L, Fink P. Training general practitioners in the treatment of functional somatic symptoms: effects on patient health in a cluster-randomised controlled trial (the Functional Illness in Primary Care study). *Psychother Psychosom* 2010;79:227–37.
60. Lamb SE, Williams MA, Williamson EM, Gates S, Withers EJ, Mt-Isa S, Ashby D, Castelnuovo E, Underwood M, Cooke MW, MINT Trial Group. Managing Injuries of the Neck Trial (MINT): a randomised controlled trial of treatments for whiplash injuries. *Health Technol Assess* 2012;16:iii–v.
61. Janse A, Wiborg JF, Bleijenberg G, Tummers M, Knoop H. The efficacy of guided self-instruction for patients with idiopathic chronic fatigue: a randomized controlled trial. *J Consult Clin Psychol* 2016;84:377–88.
62. Renkewitz F, Keiner M. How to detect publication bias in psychological research. *Z Psychol* 2019;227:261–79.
63. Petersen MW, Schröder A, Jørgensen T, Ømbøl E, Dantoft TM, Eliassen M, Thuesen BH, Fink P. The unifying diagnostic construct of bodily distress syndrome (BDS) was confirmed in the general population. *J Psychosom Res* 2020;128:109868.
64. Koch H, van Bokhoven MA, Bindels PJ, van der Weijden T, Dinant GJ, ter Riet G. The course of newly presented unexplained complaints in general practice patients: a prospective cohort study. *Fam Pract* 2009;26:455–65.
65. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457–65.
66. Augusteijn HEM, van Aert RCM, van Assen MALM. The effect of publication bias on the Q test and assessment of heterogeneity. *Psychol Methods* 2019;24:116–34.
67. Pustejovsky JE, Rodgers MA. Testing for funnel plot asymmetry of standardized mean differences. *Res Synth Methods* 2019;10:57–71.
68. Vevea JL, Woods CM. Publication bias in research synthesis: sensitivity analysis using a priori weight functions. *Psychol Methods* 2005;10:428–43.
69. Hofmann SG, Hayes SC. The future of intervention science: process-based therapy. *Clin Psychol Sci* 2019;7:37–50.
70. Weigel A, Hüsing P, Kohlmann S, Lehmann M, Shedden-Mora M, Toussaint A, Löwe B, EURONET-SOMA Group. A European research network to improve diagnosis, treatment and care for patients with persistent somatic symptoms: work report of the EURONET-SOMA conference series. *J Psychosom Res* 2017;97:136–8.
71. Löwe B, Piontek K, Daubmann A, Härter M, Wegscheider K, König H-H, Shedden-Mora M. Effectiveness of a Stepped, Collaborative, and Coordinated Health Care Network for Somatoform Disorders (Sofu-Net): a controlled cluster cohort study. *Psychosom Med* 2017;79:1016–24.
72. Holsting AF, Rask MT, Frostholm L, Rosendal M, Rask CU. Self-help interventions for young people with persistent physical symptoms: a systematic review. *J Psychosom Res* 2021;148:110553.
73. Henningsen P, Gündel H, Kop WJ, Löwe B, Martin A, Rief W, Rosmalen JGM, Schröder A, van der Feltz-Comelis C, van den Bergh O, EURONET-SOMA Group. Persistent physical symptoms as perceptual dysregulation: a neuropsychobehavioral model and its clinical implications. *Psychosom Med* 2018;80:422–31.