REVIEW ARTICLE

EJP
European Journal of Pain

Longitudinal outcome evaluations of Interdisciplinary Multimodal Pain Treatment programmes for patients with chronic primary musculoskeletal pain: A systematic review and meta-analysis

Stefan Elbers[1,2] (iD) | Harriët Wittink[1] (iD) | Sophie Konings[3] (iD) | Ulrike Kaiser[4,5] (iD) | Jos Kleijnen[6] | Jan Pool[1] (iD) | Albère Köke[2,7,8] (iD) | Rob Smeets[2,9] (iD)

[1]Research group Lifestyle & Health, Research Centre Healthy and Sustainable Living, University of Applied Sciences Utrecht, Utrecht, The Netherlands

[2]Department of Rehabilitation Medicine, Research School CAPHRI, Faculty of Health, Life Sciences and Medicine, Maastricht University, Maastricht, The Netherlands

[3]Department of Health Innovation and Technology, Fontys University of Applied Sciences, Eindhoven, The Netherlands

[4]Comprehensive Pain Center, Medical Faculty Technical University Dresden, Dresden, Germany

[5]University Hospital Carl Gustav Carus Dresden, Dresden, Germany

[6]Department of Family Medicine, Research School CAPHRI, Faculty of Health, Life Sciences and Medicine, Maastricht University, Maastricht, The Netherlands

[7]Centre of Expertise in Pain and Rehabilitation, Adelante, Maastricht, The Netherlands

[8]South University of Applied Sciences, Heerlen, The Netherlands

[9]CIR Revalidatie, Eindhoven, The Netherlands

Abstract

Background and objectives: Although Interdisciplinary Multimodal Pain Treatment (IMPT) programmes share a biopsychosocial approach to increase the wellbeing of patients with chronic pain, substantial variation in content and duration have been reported. In addition, it is unclear to what extent any favourable health outcomes are maintained over time. Therefore, our first aim was to identify and analyse the change over time of patient-related outcome measures in cohorts of patients who participated in IMPT programmes. Our second aim was to acquire insight into the heterogeneity of IMPT programmes.

Databases and data treatment: The study protocol was registered in Prospero under CRD42018076093. We searched Medline, Embase, PsycInfo and Cinahl from inception to May 2020. All study selection, data extraction and risk of bias assessments were independently performed by two researchers. Study cohorts were eligible if they included adult patients with chronic primary musculoskeletal pain for at least 3 months. We assessed the change over time, by calculating pre-post, post-follow-up and pre-follow-up contrasts for seven different patient-reported outcome domains. To explore the variability between the IMPT programmes, we summarized the patient characteristics and treatment programmes using the intervention description and replication checklist.

Results: The majority of the 72 included patient cohorts significantly improved during treatment. Importantly, this improvement was generally maintained at follow-up. In line with our expectations and with previous studies, we observed substantial methodological and statistical heterogeneity.

Correspondence
Stefan Elbers, University of Applied Sciences Utrecht, Research group Lifestyle and Health, P.O. Box 12011, 3508 AA Utrecht, The Netherlands.
Email: stefan.elbers@hu.nl

Conclusions: This study shows that participation in an IMPT programme is associated with considerable improvements in wellbeing that are generally maintained at follow-up. The current study also found substantial heterogeneity in dose and treatment content, which suggests different viewpoints on how to optimally design an IMPT programme.

Significance: The current study provides insight into the different existing approaches regarding the dose and content of IMPT programs. This analysis contributes to an increased understanding of the various approaches by which a biopsychosocial perspective on chronic pain can be translated to treatment programs. Furthermore, despite theoretical and empirical assertions regarding the difficulty to maintain newly learned health behaviors over time, the longitudinal analysis of health outcomes did not find a relapse pattern for patients who participated in IMPT programs

1 | INTRODUCTION

Interdisciplinary multimodal pain treatment (IMPT) programmes are recognized as treatment of choice for patients with chronic pain (Gatchel et al., 2014; Turk, 2003). Since the 1970s these programmes have evolved towards interventions that combine (cognitive) behavioural approaches with exercise, medical treatment and education based on a biopsychosocial model. The aim of these programmes is not to target pain itself, but to help patients to optimize daily life functioning and to increase social, physical and psychological wellbeing (Gatzounis et al., 2012; Kaiser et al., 2017; Penney & Haro, 2019). This approach is typically provided by rehabilitation centres or hospitals and requires the expertise of an interdisciplinary team of healthcare providers. Generally, these disciplines cover the biopsychosocial spectrum and continuously coordinate their treatment activities and align them to patient-specific goals.

Despite common historical roots and a biopsychosocial perspective on chronic pain, substantial variation in content, duration and outcome evaluations of IMPT programmes has been reported. For example, systematic reviews found that the total treatment duration varied between 6.4 and 196.8 h, programmes were delivered in both inpatient and outpatient settings, and pain-related disability was measured with 12 different measurement instruments (Kamper et al., 2014; Scascighini et al., 2008; Waterschoot et al., 2014). This variability not only hinders a meaningful interpretation of pooled effect sizes but it also reflects uncertainty regarding optimal dose, content and the selection of measurement instruments (Waterschoot et al., 2014).

A second problem regarding the current evidence-based IMPT programmes is that it is unclear to what extent treatment gains are maintained over time. Although RCTs often indicate a statistically significant effect compared to control interventions, post-treatment assessments still suggest a considerable impact on daily life functioning (Kamper et al., 2014). From a clinical perspective, this indicates that most patients continue to experience the burdening effect of pain after treatment. This may be problematic, as the newly learned pain management strategies are considered to be fragile and vulnerable to disruptions (e.g. unexpected exacerbations of symptoms or an unforeseen event in the personal context or nocebo's). Continuing occurrences of pain interference could prompt pre-treatment coping strategies, resulting in a declined effect over time (Carver & Scheier, 2017; Vlaeyen et al., 2016). Although this so-called 'triangular relapse pattern'—with an improvement from pre-intervention to post-intervention, followed by an unfavourable trend at follow-up—has been observed in other healthcare domains, this topic has been neglected in the field of pain rehabilitation (Brouwer et al., 2019; Morley, 2008; Opozda et al., 2016; Turk & Rudy, 1991; Wilson, 2010; Wood & Neal, 2016). To understand the impact of these programmes on patients' ability to self-regulate their wellbeing after completion of the treatment programme, it is crucial to assess the change of IMPT outcomes over time.

To acquire insight into both evidence-gaps, our first aim was to identify and analyse the change over time of key outcome measures in patients with chronic pain who participated in IMPT programmes. Therefore, the first research question is: How does the physical, psychological and social wellbeing of patients with chronic musculoskeletal pain who participated in IMPT programmes change over time? Our second aim was to explore the heterogeneity of study, patient, intervention and outcome characteristics: To what extent do cohorts vary with respect to study, patient and treatment characteristics?

2 | LITERATURE SEARCH METHODS

2.1 | Protocol and registration

The study was reported in line with the PRISMA guidelines (Moher et al., 2009, 2015) and the study protocol has been registered in PROSPERO under CRD42018076093.

2.2 | Search

We performed our search in Medline and Embase via OVID, and PsycInfo and Cinahl via EBSCOhost from inception to May 2020. The search string was developed by experienced reviewers (SE and JK) and consisted of multiple blocks that were combined with Boolean operators (see File S1). Each block included free-text words as well as specific subject headings. In addition, we searched for grey literature including unpublished studies in the Dart Europe, Open access Theses and Dissertations, NDLTD, ClinicalTrials.gov and WHO ICTRP databases. For each included study, we also performed forward (in Google Scholar) and backward reference searches.

2.3 | Eligibility criteria

Randomized controlled trials, as well as case series and cohort studies were included. Study cohorts had to include adult patients with chronic primary musculoskeletal pain for at least 3 months that was primarily perceived in musculoskeletal structures (e.g. bones, joints, muscles or related soft tissues; Treede et al., 2015, 2019). In case of mixed cohorts, at least 75% of the patients had to experience musculoskeletal pain. The criteria for IMPT programmes were based on the definition of Gatchel et al. and had to include (a) a common philosophy treatment in line with the biopsychosocial model of pain; (b) a treatment component where patients actively participated by means of tasks, training and/or exercise; (c) at least three different healthcare professionals from various disciplines that provided the interdisciplinary treatment; (d) a single facility where each patient received treatment (Gatchel et al., 2014). This last criterion excluded care-network settings, but not multicenter trials. Although structured team meetings are considered an important aspect of IMPT programmes (Kaiser et al., 2017), we did not include this as an inclusion criterion, because we expected that not all studies would explicitly report this. Our outcomes were based on the criteria developed by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials and included physical functioning, pain interference, depression, anxiety, emotional functioning, anger, self-efficacy, social functioning and pain intensity (see protocol for rationale; Dworkin et al., 2005; Turk et al., 2003). The study had to include at least one outcome that was measured at two-time points: prior to treatment and at least 12 months after the intervention was completed. Studies that focused on patients with post-surgical pain or cancer pain, as well as studies that solely included patients on the basis of specific comorbidity (e.g. depression) were excluded. Articles published in other languages than English, German or Dutch were also excluded.

2.4 | Study selection, data extraction and risk of bias

All study selection, data extraction and risk of bias assessments were independently performed by at minimum two different researchers (UK and SE for articles in German, SK, SE and MK for articles in other languages). Researchers used pre-tested forms and compared their input to reach a consensus. In case of disagreement, the study was discussed with other researchers (HW and RS) for a final decision. Study selection was performed in two rounds. In the screening round, abstracts were screened using the Rayyan software package (Ouzzani et al., 2016). Subsequently, full-text studies were assessed on all eligibility criteria.

From the extraction round onwards, we considered patient cohorts—not journal articles—as our primary unit of analysis. In case of multiple articles describing the same cohort, we combined these sources to construct a complete overview of the development over time. The first published article that met our eligibility criteria was used as the primary source and we consulted additional sources, such as protocols or follow-up studies if they contained additional relevant information. If the information sources did not contain all data items of interest, we did not contact the study authors but coded this as 'not reported' in our dataset. Our data extraction form included all items from the template for intervention description and replication (TIDieR) checklist to describe the content of the treatment programme in detail (File S2; Hoffmann et al., 2014). Risk of Bias was assessed with the Joanna Briggs Institute Checklist for Case Series, which included 10 criteria (Moola et al., 2017). A response of 'no' to any one of the items resulted in a high risk of bias, unless we found a clear indication of a limited impact of that item on the overall study outcome. The risk of bias form, including the scoring instructions, are available in the online multimedia appendix.

2.5 | Data analysis

The data extraction form included sample size (per measurement moment), age, sex, pain duration, nationality,

method of recruitment, patient eligibility criteria, exclusion criteria, study design, type of outcome measures, and outcomes for all available time points on measurement instruments of interest. If treatment intensity was expressed in days, we assumed 6 h of treatment per day. Because IMPT programmes are generally considered as a treatment of last resort, we specifically paid attention to obtaining information on attrition (Jeffery et al., 2011). We obtained pre-, post- and final-follow-up sample sizes to calculate attrition rates for post-treatment and follow-up. When a cohort presented data for two or more outcome measures within one domain, we selected the most commonly used instrument.

2.5.1 | Descriptive analysis

To investigate the heterogeneity between the included IMPT programmes, study, patient and intervention characteristics were summarized in tables. Intervention descriptions were extracted and each separate component was then classified into one of 10 possible categories. These components indicate the various means by which each IMPT programme aims to optimize daily life functioning and wellbeing. Education referred to modalities that were primarily concerned with the transfer of information from healthcare providers or experts to patients. All modalities regarding physical training, such as stretching, hydrotherapy and walking were categorized as exercise. Graded activity was only coded if the modality explicitly used the term graded activity or if the activities gradually and time-contingent increased after a baseline measurement. Modalities that described (cognitive) behavioural approaches, including problem-solving training, exposure in vivo, rational emotive therapy or ACT were classified as (cognitive) behavioural treatment. Breathing techniques, autogenic training, mindfulness, and applied relaxation techniques were classified as relaxation. Although self-management, defined as 'the intrinsically controlled ability of an active, responsible, informed and autonomous individual to live with the medical, role and emotional consequences of his chronic condition(s) in partnership with his social network and the healthcare provider(s)' (Van De Velde et al., 2019), is likely to be influenced various treatment components, some IMPT programmes included specific treatment sessions where coping with pain, setting realistic life goals and problem-solving skills were discussed. These types of sessions were classified as generic self-management skill training. Pharmacological treatment was only coded when medication was provided in response to chronic pain. Medication withdrawal procedures were coded as 'other'. Workplace visits, and ergonomic advice at the workplace were coded as workplace advice. The category body awareness included physical awareness and psychomotor exercises that aimed to improve the recognition

of bodily signals. The last category—team meetings—was only coded when the patient actively participated in the team meetings. The categories were inductively developed by first extracting and then clustering the modalities of the first search into global categories (by SE and SK). In the final dataset, these 10 categories covered more than 90% of the treatment modalities. All remaining modalities were coded as 'other'. The description of each of the modalities and the classification were registered. A similar process was performed for healthcare providers. The following professions were coded as 'physician': occupational physician, rehabilitation physician, general practitioner and not otherwise specified physician. Other physician specialists (e.g. psychiatrist, orthopaedic surgeon, anesthesiologist) who were mainly involved in consulting instead of a coordinating role were coded as 'other'. Disciplines such as clinical psychologists, general psychologists and behavioural therapists were classified as 'psychologist'. Physical therapists and physiotherapists were classified as 'physical therapist'. Social workers and social counsellors were classified as 'social worker'. Occupational therapists and nurses were classified accordingly.

To assess to what extent treatment programmes aligned their programme with individual patient characteristics and preferences (i.e. tailoring), we classified each programme into low, medium or high tailoring. We defined low tailoring as any form of personalized goal-setting, because this would allow patients to relate treatment content and progress to their personal situation. All studies received at minimum a 'low' tailoring classification because we assumed that all interdisciplinary programmes require some form of collaborative goal-setting at the start of treatment. We classified programmes as medium tailoring, when they selected or optionally provided specific treatment components based on patient-specific needs or preferences. High tailoring involved a fully personalized treatment programme, with varying duration and treatment activities and modules, based on each patient's clinical assessment.

2.5.2 | Main data analysis

In addition to pain intensity, we included seven key outcome measures as outcomes in this analysis, divided over three domains: physical health, mental health and social health. For physical health, we included physical functioning and pain interference. We extracted of the outcomes depression, anxiety, anger, and self-efficacy beliefs within the mental health domain. For social health we only included social functioning. All outcomes were defined in the study protocol. For each of these outcomes that were present within a cohort, we used the available data to calculate effect sizes for pre-post, post-follow-up and pre-follow-up contrasts. To calculate effect sizes, we used the method of

Becker et al's standardized mean change (1988), with the modifications that were suggested by Morris (2000). The model assumes that the outcomes are normally distributed at both time points, with separate means but equal variances. Furthermore, the model corrects for a pre-post within-group correlation. Because we did not have access to the original data of the included cohorts, we imputed this value (Lipsey & Wilson, 2001). For all studies, we imputed the median correlation ($r = .59$) of a meta-analysis that investigated the range of within-group correlation values in active treatment groups (Balk et al., 2012). This value is comparable to other studies that have imputed within-group meta-analyses (Clond, 2016; Roberts et al., 2017). In addition, a sensitivity analysis for the within-subject correlation is available in the multimedia appendix for all *r*-values between 0 and 1. Sample sizes lower than $n = 10$ were not included in the analysis, because this could lead to inaccurate estimates of the standardized mean gain (Morris, 2000). In some interventions, the main treatment programme was followed by follow-up treatment activities to enhance maintenance. In these situations, we considered end of treatment as the moment that the main treatment programme (ie. that covered the core of the treatment procedures) ended. Hence, follow-up meetings, booster sessions or reinforcement sessions were not considered as main treatment and could continue after post-treatment assessments. All assessments within 1 month after end of treatment were considered as a 'post' measure. We used the last available time point for the follow-up contrast. We calculated standard deviations from standard errors by multiplying them with the square root of the corresponding sample size (Higgins et al., 2019). If medians and range were provided, we used the formula of Hozo et al. (2005) to estimate the mean value and corresponding SD. For studies that presented change scores, we calculated final value mean scores and imputed the baseline standard deviation. If the latter was not available, the study was not included in the meta-analysis. For medians and interquartile ranges (IQRs), we estimated means and SDs using the assumption that the IQR width is 1.35 SD (Higgins et al., 2019). In case of missing measures of variability at follow-up, we imputed the baseline value or otherwise used the mean SD of the remaining trials that reported on that outcome. If data of the cohort was presented for different subgroups, we calculated one composite mean and SD. For data that was only presented in figures (e.g. boxplots), we measured the central tendency and measure of dispersion if the figure was of sufficient quality.

Subsequently, we summarized the effect sizes per outcome, by describing the direction of effect for each of the included cohorts over time. We a priori decided not to perform any pooling, because this was not in line with our study aims and we expected substantial heterogeneity among the included studies. To facilitate interpretation of the effect sizes, we re-expressed the median pre-post effect size on the most commonly used measurement instrument, using the weighted standard deviation of all available post-intervention scores of that instrument. To assess the statistical heterogeneity of the study outcomes we also calculated the I^2 and the Q test for each outcome domain at every time point. A statistically significant Q test rejects the hypothesis that all effect sizes are equal (Huedo-Medina et al., 2006). In addition, the I^2 index provides an indication of the proportion of variability in observed effects that is either due to between-study variability or due to within-study variability (ie. sampling error; Borenstein et al., 2017). This analysis was performed with the *R* metaphor package in RStudio (R Core Team, 2013; RStudio Team, 2020; Viechtbauer, 2010).

2.5.3 | Exploratory data analysis

To further explore the included cohorts, we developed an online multimedia appendix that contains interactive forest plots for each outcome, time series and study characteristic tables. The appendix can be accessed via https://stefanelbers.shinyapps.io/deployment/. The time series show the development over time of a measurement on a standardized scale (expressed as a percentage of the maximum score of each particular measurement instrument), as well as the raw scores and standard deviations. To standardize the scores, we obtained the distance of each mean and the unfavourable end of the scale, divided by the total distance of the scale and multiplied by 100. These plots also allow for the comparison of specific cohorts over time with respect to a particular outcome. To accommodate future updates of this systematic review, the appendix also contains contact information to encourage readers to pinpoint any inaccuracies or to suggest cohorts that have not yet been included in the current review. Finally, all data extraction files are listed in this appendix, including any comments that have been made regarding handling specific difficulties with that specific cohort (e.g. dealing with change scores, or imputing missing SDs). File S3 contains the R code that has been used for all analyses in this review as well as the deployment of the appendix. These files can also be accessed in a Github repository via: https://github.com/stefanelbers/impt.meta_analysis/tree/master/deployment.

3 | RESULTS

3.1 | Results of the search

The initial search was performed in May 2019 and was updated in May 2020 using the last date of the original search as

the beginning date for the update. In total, the search yielded 31,933 hits. After deduplication, 17,988 studies remained. The screening of title and abstracts yielded 380 hits. In the final selection round, we obtained the full-text versions and included 66 studies. 314 studies were excluded: 50 studies due to study design or publication type, 41 studies related to patient criteria, 89 studies due to intervention criteria, 89 studies because they did not include outcomes within the scope of this study or did not include a follow-up measurement of 12 months or longer, 38 studies were duplicates and seven studies due to language or inclusion of patients with specific comorbidities. Seven of the included studies did not provide the necessary data (ie. central tendency and measure of dispersion for each time point) to be included the quantitative analyses and were only included in the characteristics tables. The study flow is depicted in Figure 1.

3.2 | Characteristics of the included studies

3.2.1 | Study characteristics

We included 37 case series (i.e. longitudinal studies with prospectively collected data on one cohort of patients), 20

RCTs, four N-RCTs and five other types of study design. In total, these studies provided data on 76 cohorts with a median final follow-up measurement of 12 months (range 12–120). Pre-treatment sample sizes of the cohorts ranged from 10 to 2089, with a median of 97. After filtering out cohorts that reported no dropout in the study flow (assuming a complete case analysis) the median dropout ratio for the 55 remaining cohorts was 8.56 (range 0–42.11) at post-treatment and 16.67 (range = −3.85 to 62.17) at the final follow-up, using the posttreatment sample size as a reference. Six of the 66 studies were evaluated as low risk of bias, indicating that in the majority of the cohorts we identified at least one factor that threatened the internal validity. Statistical analysis and attrition (84.8%), unclear inclusion criteria (43.9%), incomplete inclusion of participants (60.6%), and incomplete reporting of clinical information of the participants (34.8%) were the most frequent reasons for assigning an unclear or high risk of bias. File S3 provides an overview of the risk of bias assessment.

3.2.2 | Patient characteristics

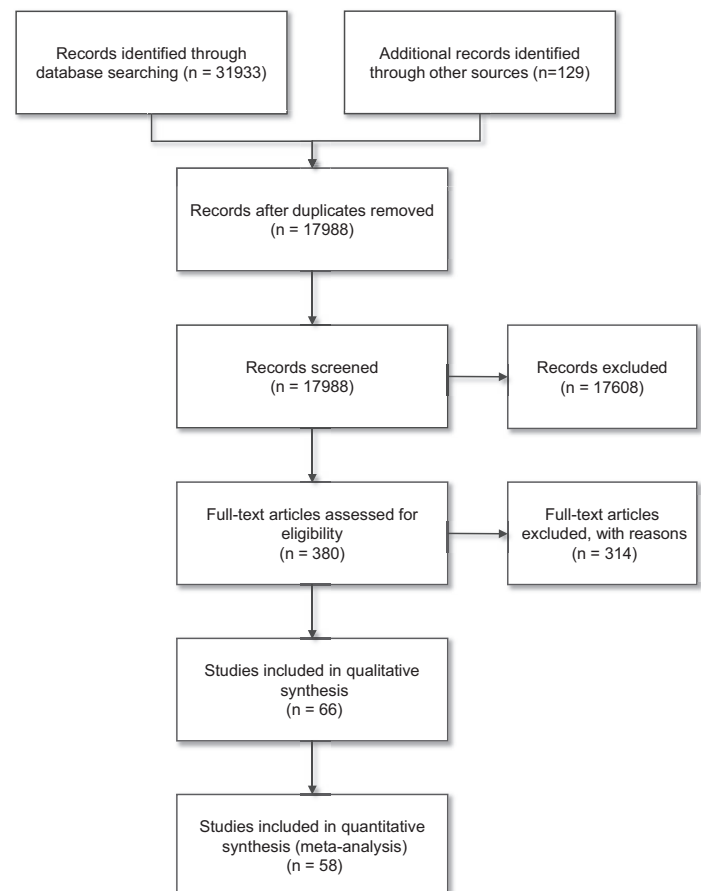The majority of the included patients were treated in European or North-American countries (94.6%). The

FIGURE 1 Flow chart of study records

mean distribution of sex was 68.7% female (SD = 18.1%) The average mean age of the cohorts was 44.353 years (SD = 4.49). The median pain duration prior to treatment, reported in 40 of the included cohorts, was 76.8 months (range = 16–217.2). Generally, patients were referred to the programme by their primary care physician or medical specialist. Table 1 provides a summary of the patient characteristics. The multimedia appendix includes a more extensive and searchable table.

3.2.3 | Intervention characteristics

The treatment aims of the included programmes often involved multiple objectives, such as increasing physical activity, return to work, or the acquisition of pain self-management skills. The median time span of the treatment duration was 5 weeks (range 1–15), with a mean intensity of 95.91 h (SD = 52.72). Twenty-two cohorts (29.73%) included, at least partly, inpatient treatment programmes and two cohorts (2.67%) solely provided treatments to individuals, whereas the other cohorts at least partially provided treatment to groups. The majority of interdisciplinary treatment was provided in secondary or tertiary care settings, such as hospitals or rehabilitation centres, with a community centre (one cohort), hotel (one cohort) and a primary care setting (three cohorts) as exceptions. Exercise (93%), education (89%), relaxation (80%), generic self-management skill training (74%), and (cognitive) behavioural treatment (70%) were included in the majority of the treatment programmes, whereas body awareness (25%), graded activity (16%), workplace advice (16%), pharmacological treatment (15%), and team meetings that included the patient (11%) were less frequently reported. The median number of these specific treatment modalities was 5 (range 2–8) per cohort. Many cohorts also included modalities that were categorized as 'other', such as assertiveness training, spinal mobilisations, group discussions, and assistance with withdrawal from pain medication. The median number of involved healthcare professionals was 4 (range 3–7), with physical therapists (97%), psychologists (93%), physicians (85%), occupational therapists (53%), nurses (34%) and social workers (32%) as respective frequencies. Other involved healthcare providers (described in 57% of the cohorts) included nutritionists, massage therapists and Qigong instructors.

Follow-up treatment sessions were described in 22 (41%) of the cohorts and mainly consisted of group refresher meetings or follow-up phone calls. Four cohorts included an extensive follow-up module where parts of the treatment programme were continued for a prolonged period (Monticone et al., 2013, 2016; Tavafian et al., 2011; Westman et al., 2006). In total, 81% of the included studies provided low tailoring, 11% medium tailoring and 8% high tailoring. Table 2 depicts a general overview of the intervention characteristics, but the full table is displayed in the multimedia appendix.

3.2.4 | Effect of time

For all outcomes, the median pre-post effect sizes show a favourable trend, indicating a positive change in health from pre to post-treatment (range = 0.38–1.94). The post to final-follow-up effect sizes vary from −0.49 to 0.15, indicating different trends. The median effect sizes from pre to final follow-up show an overall favourable change in health outcomes during the course of the study (range = 0.32 to 0.85). Table 3 shows the median effect size, range and the amount of statistical heterogeneity per contrast. The table also includes an overview of the number of cohorts that follow a particular pattern of effect over time, symbolized by different plotlines. For example, a statistically significant favourable pre-post effect that is followed by no effect from post to follow-up is represented by a positive slope, that flattens halfway.

The general trend across all outcomes indicates a statistically significant favourable effect of time in 85% of the pre-post effect sizes. This is reflected in a positive median effect size (median SMC = 0.63, range = −0.21 to 4.93). Fifteen percent of the effect sizes show no pre-post effects and there were no statistically significant unfavourable effects. For all cohorts that included a measurement at pre, post and follow-up time points, a pattern with a significant pre-post effect that is maintained at follow-up was found in 79 (51%) of the cases. Twenty-three patterns (15%) indicated a favourable pre-post effect that further improved at follow-up. A triangular relapse pattern was found in 31 (20%) of the calculated effect sizes. Two patterns (1%) showed no effect from pre to post, but a positive effect from post to follow-up and 17 outcomes (11%) did not show any effect from pre to post or from post to follow-up. Four outcomes (3%) showed no pre-post effect, but an unfavourable effect from post to follow-up. Finally, the dataset did not contain any pattern with statistically significant unfavourable pre-post results. Not all studies included a post-treatment measure, which explains why the pre to final follow-up evaluations include more effect sizes than in the pattern analysis. The effect of time from pre to final follow-up was favourable for 174 (76%) of the effect sizes. Fifty-one effect sizes (22%) did not indicate an effect and four effect sizes (2%) showed an unfavourable effect over time. In the multimedia appendix, we provided time series plots where we standardized each outcome measure on a scale from 0 to 100 (percentage of maximum score of the measurement instrument) and plotted the

TABLE 1 Patient characteristics

Author (year) Cohort	Cohort name	Study sample size	Dropout (%)	Lost to follow-up (5)	Nationality	% females	Mean age (SD)	Patient group	Mean pain duration (months)
Abbasi et al. (2012): a	Spouse-assisted multidisciplinary pain management program	36	10	0	Iran	88	45 (10)	Low back pain	107.5
Abbasi et al. (2012): b	Patient-oriented multidisciplinary pain management program	36	0	16.67	Iran	88	45 (10)	Low back pain	107.5
Beaudreuil et al. (2010)	Functional restoration program	39	0	0	France	28	42.9 (8.2)	Low back pain	28.6
Bendix et al. (1998)a	Functional restoration (A)	238	14.55	6.38	Denmark	67.4	42 (NA)	Chronic low back pain	NA
Bendix et al. (1998): b	Functional restoration (B)	238	10.87	9.76	Denmark	70.27	41 (NA)	Chronic low back pain	NA
Bendix et al. (2000)	Functional Restoration	138	13.56	5.88	Denmark	66.1	42.1 (NA)	Chronic low back pain	NA
Bergström et al. (2001)	Vocational rehabilitation program	156	NA	NA	Sweden	48	42.5 (9.5)	neck pain or back pain	38
Bergström et al. (2014)	Multimodal rehabilitation	43	9.3	12.82	Sweden	80	41 (NA)	Disabling chronic pain	NA
Bileviciute-Ljungar and Norrefalk (2014): a	Rehabilitation cohort (1998)	88	NA	NA	Sweden	84	43.2 (2.1)	Persistent musculoskeletal pain	NA
Bileviciute-Ljungar and Norrefalk (2014): b	Rehabilitation (2003)	88	NA	NA	Sweden	89.47	39.5 (1.9)	Persistent musculoskeletal pain	NA
Borys et al. (2015)	Multimodal therapy	155	0	59.35	Germany	55.6	58.3 (10.4)	Chronic low back pain	217.2
Brendbekken et al. (2016)	Multidisciplinary intervention	284	2.84	37.96	Norway	53.9	41.3 (NA)	Musculoskeletal pain	NA
Cardosa et al. (2012)	MENANG program	102	0	0	Malaysia	64.3	42.87 (9.87)	Chronic pain	66.81
de Rooij et al. (2014)	Multidisciplinary treatment programme	138	9.77	5	Netherlands	95	45.04 (10.3)	Chronic widespread pain	84
Dysvik et al. (2013)	Multidisciplinary pain management programme	117	11.11	8.65	Norway	79	45 (11.25)	Chronic non-malignant pain	89.9

(Continues)

TABLE 1 (Continued)

Author (year) Cohort	Cohort name	Study sample size	Dropout (%)	Lost to follow-up (5)	Nationality	% females	Mean age (SD)	Patient group	Mean pain duration (months)
Frost et al. (2000)	Functional restoration	129	1.63	19.83	United Kingdom	45	43 (9)	Constant pain	NA
Gantschnig et al. (2017)	BAI-REHA	30	13.33	−3.85	Swiss	43	44.83 (12.57)	ICD–10 diagnosis of chronic musculoskeletal pain	NA
Gerdle et al. (2016)	Multimodal rehabilitation program	464	NA	NA	Sweden	81.6	38.1 (10.1)	Chronic pain	83.8
Grahn et al. (2000)	Multidisciplinary rehabilitation	236	NA	NA	Sweden	82	44.3 (9.1)	Prolonged musculoskeletal disease as main diagnosis	68.4
Gustafsson et al. (2002)	Multimodal, multidisciplinary outpatient rehabilitation programme	23	4.35	22.73	Sweden	100	43.8 (10.7)	Fibromyalgia or widespread chronic pain	158.4
Hafenbrack et al. (2013): a	Berlin	681	2.86	55.51	Germany	49.2	45.81 (9.22)	Back pain	NA
Hafenbrack et al. (2013): b	Hamburg	681	3.18	55.26	Germany	55.3	44.95 (9.84)	Back Pain	NA
Haiduk et al. (2017)	Zurzacher Interdisziplinäres HWS Konzept (ZIHKo)	115	NA	NA	Switzerland	83.1	40.3 (12.3)	Chronic neck pain	17.2
Hällstam et al. (2016)	Multimodal rehabilitation	42	NA	NA	Sweden	90.5	43.6 (15.7)	Complex pain problems	NA
Hazard et al. (1989)	Functional restoration	64	1.69	31.03	USA	36	37 (4.3)	Low back pain	NA
Hildebrandt et al. (1996)	GRIP	98	1.1	6.67	Germany	48.9	41.7 (8.7)	Lumbal and unspecific back pain	NA
Huffman et al. (2019)	ICPRP	2089	19.53	62.17	USA	65.3	46.6 (13.6)	chronic non cancer pain	NA
Ibrahim et al. (2019)	Geneva MBR program	201	20.4	33.12	Switzerland	40.8	39.98 (10.06)	non-specific LBP	NA
Jensen et al. (1997): a	multimodal cognitive behavioral treatment only (regular intervention)	63	6.67	10.71	Sweden	100	43 (9)	Nonspecific spinal pain without neurological signs	51

(Continues)

TABLE 1 (Continued)

Author (year) Cohort	Cohort name	Study sample size	Dropout (%)	Lost to follow-up (5)	Nationality	% females	Mean age (SD)	Patient group	Mean pain duration (months)
Jensen et al. (1997): b	Multimodal cognitive behavioral treatment (experimental program)	63	6.06	6.45	Sweden	100	45 (8)	Nonspecific spinal pain without neurological signs	44
Kääpä et al. (2006)	Multidisciplinary rehabilitation program	64	7.81	16.95	Finland	100	46 (7.9)	Low back pain	16
Koopman et al. (2004)	Multidisciplinary occupational training program	68	5.88	20.31	Netherlands	47.1	41.3 (8.91)	Low back pain	72
Lemstra and Olszynski (2005)	Multidisciplinary rehabilitation	43	16.28	2.78	Canada	86	49.7 (9.57)	Chronic widespread pain	121.7
Letzel et al. (2019)	Multidisciplinary biopsychosocial rehabilitation	113	28.32	20.99	Germany	67.9	59.7 (10.7)	Chronic neck pain	NA
Mangels et al. (2009): a	Behavioral-medical rehabilitation	363	0	6.19	Germany	78.8	49.5 (9)	Musculoskeletal pain	NA
Mangels et al. (2009): b	Behavioral-medical rehabilitation+booster	363	0	6.72	Germany	75.6	48.3 (15.8)	Musculoskeletal pain	NA
Martin et al. (2012)	PSYMEPHY	180	10	13.89	Spain	93.46	49.07 (8.92)	Fibromyalgia	174.22
McAllister et al. (2005)	Multidisciplinary chronic pain program	276	20.65	54.79	United States	66.3	44.7 (9.7)	Chronic non-malignant pain	NA
Meng et al. (2011)	New back school	382	11.17	16	Germarny	65.2	50.2 (7.6)	Chronic low back pain	NA
Merrick and Sjölund (2009)	Interdisciplinary rehabilitation	255	1.79	−1.21	Sweden	79	39 (NA)	Disabling chronic pain	NA
Merrick et al. (2012)	Multimodal rehabilitation	296	NA	NA	Sweden	76	39.2 (9.7)	Chronic pain	62.4
Monticone et al. (2013)	Multidisciplinary intervention	90	0	0	Italy	60	48.96 (7.97)	Chronic non-specific low back pain	22.15
Monticone et al. (2016)	Multidisciplinary cognitive behavioural rehabilitation programme	150	1.33	12.16	Italy	62.67	53.2 (11.1)	Non-specific low back pain	21.7
Nagel and Korb (2009)	Multimodal therapy	351	38.96	0	Germany	59.3	44.7 (NA)	Chronic nonspecific back pain	140.4
Nicholas et al. (2014): a	Interoceptive Exposure	140	10.61	28.81	Australia	51	42.05 (12.33)	Chronic pain	67.16

TABLE 1 (Continued)

Author (year) Cohort	Cohort name	Study sample size	Dropout (%)	Lost to follow-up (5)	Nationality	% females	Mean age (SD)	Patient group	Mean pain duration (months)
Nicholas et al. (2014): b	Distraction	140	8.11	41.18	Australia	55	43.22 (11.08)	Chronic pain	77.71
Olason (2004)	Interdisciplinary pain management program	158	24.05	NA	Iceland	70.9	39.5 (NA)	Chronic pain	NA
Oslund et al. (2009)	Comprehensive outpatient pain management program	108	0	42.5	USA	70.4	55 (11.47)	Patients with various chronic pain problems not specified	110.44
Persson et al. (2012)	Musculoskeletal interdisciplinary pain rehabilitation program	813	13.53	27.6	Sweden	79	40 (9.6)	Musculoskeletal pain	49.2
Pietilä-Holmner et al. (2020)	Musculoskeletal interdisciplinary pain rehabilitation program	467	NA	NA	Sweden	85.5	43.6 (10.8)	Musculoskeletal chronic pain	NA
Reck et al. (2017)	Interdisciplinary multimodal pain therapy	71	NA	NA	Switzerland	54	43.59 (11.84)	Chronic non-specific low back pain	NA
Richardson et al. (1994)	Pain management course	109	9.17	−10.1	United Kingdom	68	45 (10)	Chronic pain	128
Roche-Leboucher et al. (2011)	Functional restoration	132	NA	NA	France	32.4	40.8 (7.4)	Non-specific low back pain	NA
Semrau et al. (2015)	PASTOR	554	5	27.44	Germany	54.1	48.9 (8)	Low back pain	NA
Silvemark et al. (2014)	Pain rehabilitation programme	164	13.64	48.87	Sweden	69.5	37.4 (9.07)	Long-term non-malignant pain	68.4
Smeets et al. (2008): a	Graded activity with problem solving training	223	5.17	5.45	Netherlands	58.6	42.52 (9.67)	Non-specific low-back pain	68.33
Smeets et al. (2008): b	Combination treatment	223	9.84	3.64	Netherlands	37.7	40.67 (10.14)	Non-specific low-back pain	56.13
Spinhoven et al. (2004)	Merged cohorts: operant-behavioral treatment +coping or group discussion (OPCO/OPDI)	148	11.97	12.62	Netherlands	63.5	39.8 (9.1)	Low back pain	117.6
Stein and Miclescu (2013)	Multidisciplinary rehabilitation	59	0	13.56	Sweden	86	48 (7.8)	Chronic non-cancer pain	NA
Steinmetz et al. (2019)	Multimodal therapy	276	9.78	56.22	Germany	57.4	53.4 (10.6)	chronic spinal pain	NA

(Continues)

TABLE 1 (Continued)

Author (year) Cohort	Cohort name	Study sample size	Dropout (%)	Lost to follow-up (5)	Nationality	% females	Mean age (SD)	Patient group	Mean pain duration (months)
Strobel et al. (1998)	Interdisciplinary group treatment	32	NA	NA	Germany	87.5	49 (7.5)	Fibromyalgia (ACR criteria)	92.4
Tavafian et al. (2011)	Group-based multidisciplinary rehabilitation program	197	4.12	25.81	Iran	73.2	44.6 (10.2)	Chronic low back pain	75.9
Thieme et al. (2003)	Operant pain therapy	61	4.76	0	Germany	100	46.61 (8.67)	Fibromyalgia	204.6
Van der Maas et al. (2015): a	Pain rehabilitation	114	42.11	21.21	Netherlands	71.1	45.4 (11.1)	Chronic musculoskeletal pain	NA
Van der Maas et al. (2015): b	Rehabilitation treatment +body awareness	114	8.16	48.89	Netherlands	91.8	38.6 (11.1)	Chronic musculoskeletal pain	NA
van Hooff et al. (2010)	Pain management programme	107	23.36	0	Netherlands	57	44.1 (8.4)	Low back pain	147.6
van Wilgen et al. (2009)	Inpatient multidisciplinary CBT programme	32	18.75	0	Netherlands	73	42 (11)	Chronic pain	96
Vendrig et al. (2000)	Multidisciplinary behavioural program for chronic pain	147	NA	NA	Netherlands	31	41.6 (8.5)	Chronic low back pain	46.3
Verkerk et al. (2011)	Functional recovery	1760	3.64	43.1	Netherlands	74.26	40.1 (10.6)	Chronic non-specific low back pain	92.4
Volker et al. (2017)	Outpatient multidisciplinary programme	165	6.67	22.08	Netherlands	87	44.1 (12.9)	Chronic musculoskeletal pain	NA
Vowles et al. (2011)	ACT	187	8.56	36.84	United Kingdom	64.2	47.3 (11.4)	Chronic pain	96
Wagner et al. (2011)	Interdisciplinary treatment program	142	0	0	Canada	63.64	49.16 (10.03)	Chronic pain	32.99
Westman et al. (2006)	Early multimodal rehabilitation	91	NA	NA	Sweden	70	41.5 (NA)	Musculoskeletal pain and disability	NA
Williams et al. (1996): a	Inpatient pain management	323	28.57	3.33	United Kingdom	53	48.7 (11.6)	Chronic pain	NA
Williams et al. (1996): b	Outpatient pain management	323	7.32	18.42	United Kingdom	49	50.4 (11.7)	Chronic pain	NA
Zhuk et al. (2018)	Multimodal pain treatment	59	NA	85.68	Germany	62.7	45.27 (10.34)	Low back pain	NA

TABLE 2 Intervention characteristics

| Cohort | Treatment modalities | Healthcare providers | | | | | | In/outpatient setting |
		phy	psy	pt	ot	nur	swo	
Abbasi et al. (2012): a	ed, bt, re, sm, oth	1	1	1	0	0	0	Out
Abbasi et al. (2012): b	ed, bt, re, sm, oth	1	1	1	0	0	0	Out
Beaudreuil et al. (2010)	ex, re, oth	1	0	1	1	0	1	Out
Bendix (1998): a	ed, ex, bt, re, oth	1	1	1	1	0	1	Out
Bendix et al. (1998): b	ed, ex, bt, re, oth	1	1	1	1	0	1	Out
Bendix et al. (2000)	ed, ex, re, sm, oth	1	1	1	1	0	1	Out
Bergström et al. (2001)	ed, ex, re, ba, oth	0	0	1	0	0	0	In
Bergström et al. (2014)	ed, ex, re, sm, ba, te, oth	1	1	1	1	0	1	Out
Bileviciute-Ljungar and Norrefalk (2014): a	ed, ex, sm, wo, oth	1	1	1	1	1	1	Out
Bileviciute-Ljungar and Norrefalk (2014): b	ed, ex, sm, wo, oth	1	1	1	1	1	1	Out
Borys et al. (2015)	ed, ex, bt, ph, ba, te	1	1	1	0	0	1	In
Brendbekken et al. (2016)	ex, sm, te	1	0	1	0	0	1	Out
Cardosa et al. (2012)	ed, ex, bt, re, sm, oth	1	1	1	0	1	0	In
de Rooij et al. (2014)	ed, ex, bt, re, sm, oth	1	1	1	1	0	1	Out
Dysvik et al. (2013)	ed, ex, sm, ba, oth	1	1	1	0	1	0	Out
Frost et al. (2000)	ed, ex, re, oth	1	1	1	0	0	0	Out
Gantschnig et al. (2017)	ed, ex, bt, re, sm, ba, wo, oth	1	1	1	1	1	1	Mix
Gerdle, Molander, Stenberg, Stalnacke, et al. (2016)	ed, bt, sm, oth	1	1	1	1	0	1	Out
Grahn et al. (2000)	ed, ex, bt, re, sm, ba, wo, oth	1	0	1	1	1	0	In
Gustafsson et al. (2002)	ed, ex, re, ba, oth	1	1	0	0	1	1	Out
Hafenbrack et al. (2013): a	ed, ex, re, sm, ph, wo	1	1	1	0	0	0	Out
Hafenbrack et al. (2013): b	ed, ex, re, sm, ph, wo	1	1	1	0	0	0	Out
Haiduk et al. (2017)	ex, bt, re, sm	1	1	1	1	0	0	In
Hållstam et al. (2016)	ex, bt, re, sm, oth	1	1	1	0	1	0	Out
Hazard et al. (1989)	ed, ex, bt, re, sm, oth	0	1	1	1	0	0	Out
Hildebrandt et al. (1996)	ed, ex, bt, re, sm, wo	1	1	1	0	1	1	Out
Huffman et al. (2019)	ed, ex, bt, oth	0	1	1	1	1	0	Out
Ibrahim et al. (2019)	ed, ex, bt, re, sm, oth	1	1	1	1	0	0	Out
Jensen et al. (1997): a	ed, ex, bt, re, sm, oth	1	1	1	0	1	0	In
Jensen et al. (1997): b	ed, ex, bt, re, sm, oth	1	1	1	0	1	0	In
Kääpä et al. (2006)	ed, ex, bt, re, ph, wo	1	1	1	1	0	0	Out
Koopman et al. (2004)	ex, ga, bt, re, oth	1	1	1	0	0	0	Out
Lemstra and Olszynski (2005)	ed, ex, re, sm, oth	1	1	1	0	0	0	Out
Letzel et al. (2019)	ed, ex, bt, re, sm, oth	1	1	1	1	0	0	Out
Mangels et al. (2009): a	ed, ex, re, sm, ph	0	1	1	1	0	0	In
Mangels et al. (2009): b	ed, ex, re, sm, ph	0	1	1	1	0	0	In
Martín et al. (2012)	ed, ex, bt, re, sm, ph	1	1	1	0	0	0	Out
McAllister et al. (2005)	ex, re, sm, oth	1	0	1	0	1	0	Mix
Meng et al. (2011)	ed, ex, sm	0	1	1	0	0	0	In
Merrick and Sjölund (2009)	ed, ex, bt, re, sm, ba, oth	1	1	1	1	0	1	Out

Type of contact	Group size	Location	Time span (wks)	Duration (h)	Level of tailoring	Follow-up sessions
Mixed	6	Pain clinic	7	15	Low	No
Mixed	6	Pain clinic	7	16	Low	No
Group	2–6	Hospital	5	138	Low	No
Group	6–8	Pain clinic	3	135	Low	No
Group	6–8	Pain clinic	3	135	Low	No
NA	6–8	Pain clinic	3	135	Low	No
Group	14	Rehab center	4	160	Medium	No
Group	10–12	Hospital	5	NA	Low	No
Mixed	8	Hospital	8	143	Low	No
Mixed	8	Hospital	8	143	Low	No
Mixed	NA	Hospital	3	168	Low	No
Individual	NA	Hospital	NA	8	Low	yes
Group	NA	Hospital	2	60	Low	No
Mixed	NA	Rehab center	7	49	High	Yes
Group	8–12	Hospital	8	45	Low	Yes
Group	max 5	NA	3	98	Low	Yes
Mixed	NA	Hospital	12	108	Medium	No
Group	6–9	Hospital	7	140	Medium	No
Mixed	NA	Rehab center	4	120	High	Yes
Mixed	7–8	Hospital	12	96	Low	Yes
Group	NA	Pain clinic	4	120	Low	No
Group	NA	Pain clinic	4	120	Low	No
Mixed	NA	Hospital	4	108	Low	No
Mixed	NA	Hospital	13	65	Low	No
Mixed	NA	Rehab center	3	53	Low	No
Group	8–10	Pain clinic	11	207	Medium	No
Mixed	NA	Rehab center	3.5	166	Low	Yes
Group	4–6	Hospital	4	100	Low	Yes
Group	NA	Pain clinic	5	NA	Low	Yes
Group	NA	Pain clinic	5	NA	Low	Yes
Mixed	6–8	Rehab center	8	70	Low	No
Mixed	6–10	Rehab center	12	216	Low	No
Mixed	NA	Community center	6	33	Low	No
Group	5–10	Hospital	3	44	Low	No
Mixed	10–12	Hospital	NA	167	Low	No
Mixed	10–12	Hospital	NA	167	Low	Yes
Group	12	Hospital	6	21	Low	No
Group	NA	NA	4	80	Low	Yes
Group	7–15	Hospital	3	163	Low	No
Mixed	NA	Hospital	5	67	Low	No

T A B L E 2 (Continued)

Cohort	Treatment modalities	Healthcare providers						In/outpatient setting
		phy	psy	pt	ot	nur	swo	
Merrick et al. (2012)	ed, ex, bt, re, sm, ba, te, oth	1	1	1	1	0	1	Out
Monticone et al. (2013)	ed, ex, bt, sm, oth	1	1	1	0	0	0	Out
Monticone et al. (2016)	ed, ex, bt, re, sm, oth	1	1	1	0	0	0	Out
Nagel and Korb (2009)	ed, ex, bt, re	1	1	1	0	1	0	Out
Nicholas et al. (2014): a	ed, ex, bt, sm, oth	1	1	1	0	1	0	Out
Nicholas et al. (2014): b	ed, ex, bt, re, sm	1	1	1	0	1	0	Out
Olason (2004)	ed, ex, bt, re, sm, ph, ba, te, oth	1	1	1	1	1	1	In
Oslund et al. (2009)	ed, ex, bt, re, sm	1	1	1	1	0	0	Out
Persson et al. (2012)	ed, ex, bt, re, sm, ba, wo, te, oth	1	1	1	1	1	1	Out
Pietilä-Holmner et al. (2020)	ed, ex, bt, re, sm	1	1	1	1	0	1	Out
Reck et al. (2017)	ed, ex, bt, re, sm, wo	1	1	1	1	0	0	Out
Richardson et al. (1994)	ed, ex, ga, bt, re, sm, oth	0	1	1	1	1	0	Mix
Roche-Leboucher et al. (2011)	ex, re, wo, te, oth	1	1	0	1	0	0	Out
Semrau et al. (2015)	ed, ex, re, sm, oth	1	1	1	0	0	1	In
Silvemark et al. (2014)	ed, ex, bt, re, sm, ba	1	1	1	1	1	1	Out
Smeets et al. (2008): a	ed, ga, bt, oth	1	1	1	1	0	1	Out
Smeets et al. (2008): b	ed, ex, ga, bt, oth	1	1	1	1	0	1	Out
Spinhoven et al. (2004)	ed, ga, bt, sm, oth	0	1	1	1	0	0	Mix
Stein and Miclescu (2013)	ed, ex, bt, re, sm, ba, te, oth	1	1	1	1	0	0	Out
Steinmetz et al. (2019)	ex, re, ph, oth	1	1	1	0	0	0	In
Strobel et al. (1998)	ed, ex, ga, re, sm, ba, oth	1	1	1	0	1	0	In
Tavafian et al. (2011)	ed, ex, bt, re, sm, ph, oth	0	1	1	0	0	0	Out
Thieme et al. (2003)	ed, ex, bt, sm	1	1	1	0	1	0	In
Van der Maas et al. (2015): a	ed, ex, ga, bt, re, oth	1	1	1	1	0	0	Out
Van der Maas et al. (2015): b	ed, ex, ga, bt, re, ba, oth	1	1	1	1	0	0	Out
van Hooff et al. (2010)	ed, ex, bt, re, sm	0	1	1	1	0	0	In
van Wilgen et al. (2009)	ed, ex, ga, bt, re, sm, oth	1	1	1	0	0	0	Mix
Vendrig et al. (2000)	ed, ex, ga, ph, wo	0	1	1	0	0	0	Out
Verkerk et al. (2011)	ed, ex, re, sm, ba	1	1	1	0	0	0	Out
Volker et al. (2017)	ed, ex, bt, re	1	1	1	1	0	1	Out
Vowles et al. (2011)	ed, ex, bt, re, sm, ba	1	1	1	1	1	0	In
Wagner et al. (2011)	ed, ex, bt, re, sm	1	1	1	1	1	0	Out
Westman et al. (2006)	ed, ex, bt, re, ba, oth	1	1	1	0	0	0	Out
Williams et al. (1996): a	ed, ex, ga, bt, re, pm, oth	0	1	1	1	1	0	In
Williams et al. (1996): b	ed, ex, ga, bt, re, pm, oth	0	1	1	1	1	0	Out
Zhuk et al. (2018)	ed, ex, bt, re, sm, ba, oth	1	1	1	0	1	0	Out

Abbreviations: ba, body awareness therapy; bt, (cognitive) behavioral therapy; ed, education; ex, exercise; ga, graded activity; nur, nurse; ot, occupational therapist; oth, other type of treatment modality; ph, pharmacological therapy; phy, physician; psy, psychologist; pt, physiotherapist; re, relaxation; sm, self-management skills; swo, social worker; te, team meetings; wo, workplace advice.

Type of contact	Group size	Location	Time span (wks)	Duration (h)	Level of tailoring	Follow-up sessions
Group	6–8	Hospital	4	65	Medium	No
Individual	NA	Rehab center	5	15	Low	Yes
Group	5	Rehab center	5	15	Low	Yes
Group	8–9	Pain clinic	3.5	89	Low	No
Group	8–10	Hospital	3	115	Low	No
Group	8–10	Hospital	3	115	Low	No
Mixed	NA	Rehab center	7	NA	Medium	No
Mixed	NA	University	4	120	Low	No
Group	9	Hospital	5	126	Low	Yes
Mixed	NA	Primary care centers	8	20	Low	No
Group	NA	Pain clinic	1	45	Low	Yes
Group	5	Hospital	NA	120	Low	No
Group	6–8	Rehab center	5	150	Low	No
Mixed	NA	Rehab center	3	48	Low	No
Group	6–9	Hospital	5	175	Low	No
Group	Max 4	Rehab center	10	24	Low	No
Group	Max 4	Rehab center	10	77	Low	No
Mixed	NA	Rehab center	8	150	Low	No
Group	6–8	Primary care setting	6	90	Low	No
NA	NA	Hospital	2	15	High	No
Group	6	Rehab center	5	150	Low	No
Group	NA	University	1	9	Low	Yes
Group	5–7	Hospital	5	75	Low	No
Group	4–6	Rehab center	12	94	Low	Yes
Group	4–6	Rehab center	12	109	Low	Yes
Group	NA	Hotel facility	2	100	Low	No
Mixed	NA	Hospital	7	NA	High	No
Group	NA	Rehab center	4	NA	Low	No
Group	6	Rehab center	9	48	Low	Yes
Mixed	NA	Rehab center	15	NA	Low	No
Group	NA	Hospital	3.5	114	Low	No
Mixed	NA	Pain clinic	6	123	High	No
Group	8–10	Primary care setting	8	140	Low	Yes
Group	10	Hospital	4	108	Low	No
Group	10	Hospital	8	28	Low	No
Group	NA	Pain clinic	3	NA	Low	No

development of each cohort over time using all available data points.

3.2.5 | Heterogeneity in outcomes

The Cochrane's *Q*-tests for all outcomes at each of the three contrasts were significant, except for self-efficacy at post-follow-up. For all other contrasts, this indicates that the null hypotheses that these studies are evaluating the same effect were rejected (Higgins & Thompson, 2002). In addition, only except for self-efficacy at post-follow ($I^2 = 0\%$), all values were considerably high, with the majority of the values over 90%. These analyses support our decision to refrain from pooling the effect sizes. Rather, multiple different patient, study or interventions factors may account for this variability. The self-efficacy post-follow-up contrasts indicate a stable maintenance pattern across studies.

4 | DISCUSSION

4.1 | Summary of findings

Our first objective was to investigate the development over time of patients who participated in IMPT programmes. The results indicate that the majority of the patient cohorts significantly improved from pre to post-treatment. Importantly, this was mostly maintained at final follow-up, which is in contrast to typical triangular relapse patterns that have been observed in other health behaviour change efforts (Wood & Neal, 2016). Although the results indicate that pre-post effects of IMPT are generally maintained over time, the possibility of relapse for individual patients should not be neglected. Closer inspection of the distribution of individual cohort data, such as the post to follow-up physical functioning data of Silvemark et al. (2014) (SMD = 0.06, 95% CI = −0.15 to 0.28), reveals that 47% of the patients show a decrease in physical functioning, assuming normally distributed data and a pre-test–post-test correlation of $r = 0.59$. To increase the accuracy of these rudimentary estimates, publishing the datasets along with the study, would allow for more detailed analyses on patient relapse across studies. This is especially relevant when taking into account that IMPT programmes are often considered as treatment of last resort (Jeffery et al., 2011).

Our second objective was to explore the study, patient, intervention and outcome heterogeneity of the included cohorts. In line with our expectations and with previous studies, we observed substantial methodological and statistical heterogeneity despite overlapping theoretical foundations, such as the biopsychosocial model (Geneen et al., 2017; Guzmán et al., 2001; Waterschoot et al., 2014). This heterogeneity can be explained by different policies, cultures, resources, and research traditions that have been influencing these treatment programmes over time (Kaiser et al., 2017, 2018). However, to our knowledge, the current study is the first attempt to extract and categorize the individual treatment modalities of IMPT programmes to assess the treatment content heterogeneity in these programmes. Despite a common heritage, the results of this assessment indicate that the included interventions do not share an equal underlying effect. Rather, interventions generally consist of a unique collection of multiple different action mechanisms that are generally not explicitly described.

4.2 | Strengths and limitations

We encountered several problems regarding the interpretation of the study data. First, for the majority of the cohorts, we identified a risk of bias, which negatively influences the validity of our results. Especially the study attrition rates, indicating that a substantial minority dropped out of the programme or discontinued participation, introduce significant non-response bias. Furthermore, incomplete reporting of the intervention and its outcomes remains an issue. To increase accuracy of reporting as well as improved understanding of how a particular IMPT programme may benefit patients, we restate the recommendation by Williams et al. (2012) to provide a clear rationale for that particular set of treatment components and to test this by including process measures (eg. Nicholas et al., 2014), instead of generally referring to a biopsychosocial approach. A practical tool that supports clear reporting is the TIDieR checklist, which includes clear guidance for reporting study rationale and action mechanisms (Hoffmann et al., 2014). Beyond the investigation of treatment benefits and harms of IMPT a standardized level of comparison, e.g. a core outcome set (Williamson et al., 2017), is required for harmonisation of outcome assessment and supporting detailed meta-analyses. Heterogeneous outcome assessment in the context of IMPT has been consistently reported (Deckert et al., 2016; Kamper et al., 2014; Waterschoot et al., 2014). Involving the patient perspective in defining helpful treatment approaches is generally recommended for such actions (Williamson et al., 2017). For IMPT an international initiative has developed a core outcome set for a domain set of outcomes, comprising the biopsychosocial impact of chronic pain in patients undergoing IMPT (Kaiser et al., 2018). Implementing such recommendations would help also to reduce reporting bias and

TABLE 3 Distribution of effect sizes of pre-post, post-follow up and pre-follow-up contrasts for each outcome domain

Outcome	Pre-post			Post-final follow-up			Pre-post-final follow-up Pattern over time[b]		Pre-final follow-up				Direction of effect[c]
	N[a]	Mdn ES (range)	Q test/I^2	N[a]	Mdn ES (range)	Q test/I^2	N[a]		N[a]	Mdn ES (range)	Q test/I^2	N[a]	
Pain interference	36	0.72 (−0.21 to 4.93) RMDQ²: 3.21	Q(df=35) = 1463.38*** I^2 = 99%	36	0.01 (−0.59 to 3.01)	Q(df=35) = 4561.80*** I^2 = 95%	7 \ / 22 / 4 < 3		53	0.53 (−0.40 to 11.46)	Q(df=52) = 1300.71*** I^2 = 99%	42 10 1	+ +/− −
Physical function	13	0.50 (0.04–3.26) FFbH²: 8.99	Q(df=12) = 225.56*** I^2 = 99%	13	0.00 (−0.23 to 0.45)	Q(df=12) = 32.64** I^2 = 68%	2 8 3		20	0.43 (0.08–3.59)	Q(df=19) = 290.58*** I^2 = 98%	16 4	+ +/−
Depression	31	0.71 (0.10–3.19) BDI²: 5.57	Q(df=30) = 889.89*** I^2 = 95%	31	−0.17 (−1.29 to 0.33)	Q(df=30) = 202.99*** I^2 = 88%	1 15 12 2 1		41	0.42 (−0.03 to 1.29)	Q(df=40) = 467.63*** I^2 = 91%	31 10	+ +/−
Anxiety	12	0.45 (0.11–3.77) HADS-A²: 1.94	Q(df=11) = 293.37*** I^2 = 99%	12	−0.04 (−1.61 to 0.73)	Q(df=11) = 66.07*** I^2 = 97%	6 4 1 1		19	0.32 (0.02–1.18)	Q(df=18) = 349.62*** I^2 = 92%	14 5	+ +/−
Emotional functioning	6	0.88 (0.48–2.99) SF36 MH²: 10.63	Q(df=5) = 111.42*** I^2 = 98%	6	−0.2 (−0.84 to 0.66)	Q(df=5) = 55.96*** I^2 = 93%	2 1 3		11	0.62 (−0.15 to 3.48)	Q(df=10) = 202.74*** I^2 = 99%	10 1	+ +/−
Anger	1	1.944 SCL-90h²: 3.88	NA	1	−0.49[d]	NA	1		1	0.49[d]	NA	1	+
Self-efficacy	9	0.88 (0.10–1.21) PSEQ²: 10.54	Q(df=8) = 83.21*** I^2 = 87.38%	9	0.06 (−0.21 to 0.25)	Q(df=8) = 7.51 I^2 = 0%	7 2		11	0.85 (0.16–1.32)	Q(df=10) = 83.5*** I^2 =84%	10 1	+ +/−

(Continues)

TABLE 3 (Continued)

Outcome	Pre-post			Post-final follow-up			Pre-post-final follow-up	Pre-final follow-up			Direction of effect[c]	
	N^a	Mdn ES (range)	Q test/I^2	N^a	Mdn ES (range)	Q test/I^2	Pattern over time[b] N^a	N^a	Mdn ES (range)	Q test/I^2	N^a	
Social functioning	5	0.38 (−0.08 to 2.37) SF36srf²: 6.39	Q (df = 4) = 126.9*** I^2 = 98%	5	0.15 (−0.27 to 0.42)	Q (df = 4) = 17.6** I^2 = 77%	╱ 1 ╱ 1 ⟨ 1 ⟩ 1 — 1	8	0.71 (0.02– 2.39)	Q (df = 7) = 129.47*** I^2 = 97%	6 2	+ +/−
Pain intensity	38	0.63 (−0.08 to 4.39) VAS²: 14.29	Q (df = 37) = 853.42*** I^2 = 99%	38	0.04 (−1.39 to 1.09)	Q (df = 37) = 761.98*** I^2 = 96%	╲ 10 ╱ 14 ⟨ 6 — 5 ╱ 3	55	0.45 (−0.31 to 4.96)	Q (df = 54) = 1481.91*** I^2 = 98%	35 17 3	+ +/− −

Abbreviations: BDI, Beck Depression Inventory; *df*, degrees of freedom; ES, effect size; FFbH, Funktionsfragebogen Hannover; HADS-A, Hospital Anxiety and Depression Scale, subscale Anxiety; Mdn, median; PSEQ, Pain Self-efficacy Questionnaire; RMDQ, Roland-Morris Disability Questionnaire; SCL-90h, Symptom Checklist 90, subscale hostility; SF-36 MH, Short Form 36, subscale mental health; SF-36: SF, Short Form 36, subscale social functioning.

[a]N represents the number of cohorts.

[b]The re-expression of the median pre-post effect size on a familiar instrument was obtained by multiplying the effect size by the weighted standard deviation of all post-intervention scores from that instrument.

[c]The pattern reflects the development of an outcome over three time points (pre, post-final follow-up). An upward slope between two points indicates a statistically significant favorable effect (i.e. the lower limit of the 95% CI of the ES > 0). A straight line represents no effect (i.e. the lower limit of the 95%CI of the ES < 0 and the upper limit > 0); a downward slope represents a statistically significant unfavorable effect (i.e. the upper limit of the 95% CI of the ES < 0).

[d]Direction of effect: + represents a statistically significant effect from pre to final follow-up; +/− represents no effect; − represents a statistically significant unfavorable effect from pre to final follow-up. No range, only one study included.

*p > .05.

**p > .01.

***p > .001.

enhance reporting quality of clinical trials (Williamson et al., 2017). On a smaller scale, the successful implementation of similar initiatives has resulted in improved collaboration between healthcare services and a homogeneous dataset (Tardif et al., 2017). It should be noted that such initiatives are either shaped by national requirements and resources, commonly organized in national registries, or aim for international application in clinical trials (Kaiser et al., 2018). An important future challenge is to harmonize these approaches in order to achieve results in both objectives. Second, the categorisation of treatment modalities is likely to contain erroneous interpretations, either due to incomplete reporting or to misinterpretation during the data extraction. Moreover, it is important to realize that the categories for the treatment modalities, the classification of tailoring and what constitutes as an IMPT programme remain arbitrary and leave room for discussion and further refinement in future studies. To provide transparency regarding our procedures and choices, we published all data extraction forms in the online multimedia appendix. Third, the calculation of pre-post effect sizes in meta-analyses is under debate (Cuijpers et al., 2017; Kösters, 2017). We realize that pre-post effects should not be considered as a valid method for demonstrating a treatment effect. However, the current analysis does provide an indication of the change over time of patients' wellbeing on several key outcomes, which is considered particularly useful from a clinician's perspective (Kösters, 2017). Of notice is that the suggested overestimation of effect size has not been observed in large comparison studies (Anglemyer et al., 2014; Benson & Hartz, 2000; Concato et al., 2000). On the contrary, less heterogeneity was found in observational studies compared to RCTs. A possible explanation was that cohorts within case series potentially better represent the population at risk and tailor treatment to specific patients, compared to RCTs with specific inclusion criteria and standardized protocols.

4.3 | Future directions

An opportunity to increase the lifespan and relevance of this systematic review is to develop this study into a living systematic review. The main characteristic of this type of review is that it will be continuously updated when new evidence becomes available (Elliott et al., 2017). In addition, living systematic reviews often include an online platform where datasets and data analysis syntaxes are publicly available, which may decrease duplicate work (Thomas et al., 2017). This helps to decrease the evidence to practice gap, but also to stimulate collaboration and data-sharing (Elliott et al., 2014). Data validation

by authors of the included studies, improved analyses techniques as well as semi-automated search and data extraction procedures are among the possibilities of such an initiative (Bannach-Brown et al., 2019; Thomas et al., 2017). The current multimedia appendix has been developed to accommodate future updates, which will facilitate this transition.

5 | CONCLUSIONS

In the past five decades, pain management programmes evolved from attempts to coordinate various disciplines in managing chronic pain to comprehensive interdisciplinary multimodal interventions that help patients to optimize their daily life functioning and their overall wellbeing. This study shows that participation in an IMPT programme is associated with considerable improvements in physical and psychological wellbeing that are generally maintained at follow-up. The current study also revealed that despite common roots these programmes show substantial heterogeneity with respect to dose and treatment content, which suggests different viewpoints on how to optimally design an IMPT intervention. To discuss these differences and learn from this variability, we recommend to improve the precision of describing the intervention rationale and to test the proposed mechanisms by which the intervention is expected to benefit the patient. Finally, we believe that regular updates of this review may support the critical monitoring of future developments of IMPT programmes, the possibility to correct for data extraction errors and the comparison of different treatment approaches. A living systematic review approach provides the potential to accommodate this.

CONFLICTS OF INTEREST
There is no conflict of interest.

AUTHOR CONTRIBUTIONS
Stefan Elbers, Ulrike Kaiser, Harriët Wittink and Rob Smeets conceived of the presented idea. Jos Kleijnen and Stefan Elbers developed the search strategy. Stefan Elbers, SK, Ulrike Kaiser and Jan Pool screened the studies and extracted the data. Harriët Wittink and Rob Smeets supervised the study. Stefan Elbers and Sophie Konings

developed the online multimedia appendix. All authors discussed the results and contributed to the final version of the manuscript.

ORCID

Stefan Elbers https://orcid.org/0000-0002-8928-4307
Harriët Wittink https://orcid.org/0000-0001-5592-2879
Sophie Konings https://orcid.org/0000-0002-9718-0583
Ulrike Kaiser https://orcid.org/0000-0002-9701-9028
Jan Pool https://orcid.org/0000-0002-9240-4488
Albère Köke https://orcid.org/0000-0002-4546-4576
Rob Smeets https://orcid.org/0000-0002-9503-366X

REFERENCES

Abbasi, M., Dehghani, M., Keefe, F. J., Jafari, H., Behtash, H., & Shams, J. (2012). Spouse-assisted training in pain coping skills and the outcome of multidisciplinary pain management for chronic low back pain treatment: A 1-year randomized controlled trial. *European Journal of Pain*, *16*(7), 1033–1043. https://doi.org/10.1002/j.1532-2149.2011.00097.x

Anglemyer, A., Horvath, H. T., & Bero, L. (2014). Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database of Systematic Reviews*, *2014*(4). https://doi.org/10.1002/14651858.MR000034.pub2

Balk, E. M., Earley, A., Patel, K., Trikalinos, T. A., & Dahabreh, I. J. (2012). In *Empirical assessment of within-arm correlation imputation in trials of continuous outcomes (methods research report no. AHRQ publication no. 12 (13)-EHC141-EF)*. Agency for Healthcare Research and Quality. Retrieved April 2, 2020, from https://www.ncbi.nlm.nih.gov/books/NBK115797/pdf/Bookshelf_NBK115797.pdf

Bannach-Brown, A., Przybyła, P., Thomas, J., Rice, A. S. C., Ananiadou, S., Liao, J., & Macleod, M. R. (2019). Machine learning algorithms for systematic review: Reducing workload in a preclinical review of animal studies and reducing human screening error. *Systematic Reviews*, *8*(1), 23. https://doi.org/10.1186/s13643-019-0942-7

Beaudreuil, J., Kone, H., Lasbleiz, S., Vicaut, E., Richette, P., Cohen-Solal, M., Lioté, F., de Vernejoul, M. C., Nizard, R., Yelnik, A., Bardin, T., & Orcel, P. (2010). Efficacy of a functional restoration program for chronic low back pain: Prospective 1-year study. *Joint Bone Spine*, *77*(5), 435–439. https://doi.org/10.1016/j.jbspin.2010.03.003

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*(2), 257–278. https://doi.org/10.1111/j.2044-8317.1988.tb00901.x

Bendix, A. F., Bendix, T., Labriola, M., & Boekgaard, P. (1998). Functional restoration for chronic low back pain. *Spine*, *23*(6), 717–725. https://doi.org/10.1097/00007632-199803150-00013

Bendix, T., Bendix, A., Labriola, M., Hæstrup, C., & Ebbehøj, N. (2000). Functional restoration versus outpatient physical training in chronic low back pain. *Spine*, *25*(19), 2494–2500. https://doi.org/10.1097/00007632-200010010-00012

Benson, K., & Hartz, A. J. (2000). A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine*, *342*(25), 1878–1886. https://doi.org/10.1056/nejm200006223422506

Bergström, G., Jensen, I. B., Bodin, L., Linton, S. J., & Nygren, A. L. (2001). The impact of psychologically different patient groups on outcome after a vocational rehabilitation program for long-term spinal pain patients. *Pain*, *93*(3), 229–237. https://doi.org/10.1016/s0304-3959(01)00320-7

Bergström, M., Ejelöv, M., Mattsson, M., Stålnacke, B.-M. (2014). One-year follow-up of body awareness and perceived health after participating in a multimodal pain rehabilitation programme—A pilot study. *European Journal of Physiotherapy*, *16*(4), 246–254. https://doi.org/10.3109/21679169.2014.935802

Bileviciute-Ljungar, I., & Norrefalk, J. R. (2014). Beneficial long-term effects of multiprofessional assessment vs. rehabilitation program in patients with musculoskeletal pain. *Pain Practice*, *14*(3), 228–236. https://doi.org/10.1111/papr.12053

Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I(2) is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Borys, C., Lutz, J., Strauss, B., & Altmann, U. (2015). Effectiveness of a multimodal therapy for patients with chronic low back pain regarding pre-admission healthcare utilization. *PLoS One*, *10*(11), e0143139. https://doi.org/10.1371/journal.pone.0143139

Brendbekken, R., Harris, A., Ursin, H., Eriksen, H. R., & Tangen, T. (2016). Multidisciplinary intervention in patients with musculoskeletal pain: A randomized clinical trial. *International Journal of Behavioral Medicine*, *23*(1), 1–11. https://doi.org/10.1007/s12529-015-9486-y

Brouwer, M. E., Williams, A. D., Kennis, M., Fu, Z., Klein, N. S., Cuijpers, P., & Bockting, C. L. H. (2019). Psychological theories of depressive relapse and recurrence: A systematic review and meta-analysis of prospective studies. *Clin Psychol Rev*, *74*, 101773. https://doi.org/10.1016/j.cpr.2019.101773

Cardosa, M., Osman, Z. J., Nicholas, M., Tonkin, L., Williams, A., Abd Aziz, K., Mohd Ali, R., & Dahari, N. M. (2012). Self-management of chronic pain in Malaysian patients: Effectiveness trial with 1-year follow-up. *Translational Behavioral Medicine*, *2*(1), 30–37. https://doi.org/10.1007/s13142-011-0095-2

Carver, C. S., & Scheier, M. F. (2017). *Self-regulatory functions supporting motivated action* (Vol. *4*). Elsevier.

Clond, M. (2016). Emotional freedom techniques for anxiety: A systematic review with meta-analysis. *The Journal of Nervous and Mental Disease*, *204*(5), 388–395. https://doi.org/10.1097/nmd.0000000000000483

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*, *342*(25), 1887–1892. https://doi.org/10.1056/nejm200006223422507

Cuijpers, P., Weitz, E., Cristea, I. A., & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences*, *26*(4), 364–368. https://doi.org/10.1017/s2045796016000809

de Rooij, A., de Boer, M. R., van der Leeden, M., Roorda, L. D., Steultjens, M. P., & Dekker, J. (2014). Cognitive mechanisms of change in multidisciplinary treatment of patients with chronic widespread pain: A prospective cohort study. *Journal of Rehabilitation Medicine*, *46*(2), 173–180. https://doi.org/10.2340/16501977-1252

Deckert, S., Kaiser, U., Kopkow, C., Trautmann, F., Sabatowski, R., & Schmitt, J. (2016). A systematic review of the outcomes reported in multimodal pain therapy for chronic pain. *European Journal of Pain*, *20*(1), 51–63. https://doi.org/10.1002/ejp.721

Dworkin, R. H., Turk, D. C., Farrar, J. T., Haythornthwaite, J. A., Jensen, M. P., Katz, N. P., Kerns, R. D., Stucki, G., Allen, R. R., Bellamy, N., Carr, D. B., Chandler, J., Cowan, P., Dionne, R., Galer, B. S., Hertz, S., Jadad, A. R., Kramer, L. D., Manning, D. C., ... Witter, J. (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*, *113*(1–2), 9–19. https://doi.org/10.1016/j.pain.2004.09.012

Dysvik, E., Kvaløy, J. T., & Furnes, B. (2013). Evaluating physical functioning as part of a Cognitive Behavioural Therapy approach in treatment of people suffering from chronic pain. *Journal of Clinical Nursing*, *22*(5–6), 806–816. https://doi.org/10.1111/j.1365-2702.2011.04040.x

Elliott, J. H., Synnot, A., Turner, T., Simmonds, M., Akl, E. A., McDonald, S., Salanti, G., Meerpohl, J., MacLehose, H., Hilton, J., Tovey, D., Shemilt, I., Thomas, J., Agoritsas, T., Hilton, J., Perron, C., Akl, E., Hodder, R., Pestridge, C., ... Pearson, L. (2017). Living systematic review: 1. Introduction-the why, what, when, and how. *Journal of Clinical Epidemiology*, *91*, 23–30. https://doi.org/10.1016/j.jclinepi.2017.08.010

Elliott, J. H., Turner, T., Clavisi, O., Thomas, J., Higgins, J. P., Mavergames, C., & Gruen, R. L. (2014). Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Medicine*, *11*(2), e1001603. https://doi.org/10.1371/journal.pmed.1001603

Frost, H., Lamb, S., & Shackleton, C. J. P. (2000). A functional restoration programme for chronic low back pain: A prospective outcome study. *Physiotherapy*, *86*(6), 285–293. https://doi.org/10.1016/S0031-9406(05)61002-4

Gantschnig, B. E., Heigl, F., Widmer Leu, C., Bütikofer, L., Reichenbach, S., & Villiger, P. M. (2017). Effectiveness of the Bern Ambulatory Interprofessional Rehabilitation (BAI-Reha) programme for patients with chronic musculoskeletal pain: A cohort study. *Swiss Medical Weekly*, *147*, w14433. https://doi.org/10.4414/smw.2017.14433

Gatchel, R. J., McGeary, D. D., McGeary, C. A., & Lippe, B. (2014). Interdisciplinary chronic pain management: past, present, and future. *American Psychologist*, *69*(2), 119–130. https://doi.org/10.1037/a0035514

Gatzounis, R., Schrooten, M. G., Crombez, G., & Vlaeyen, J. W. (2012). Operant learning theory in pain and chronic pain rehabilitation. *Current Pain and Headache Reports*, *16*(2), 117–126. https://doi.org/10.1007/s11916-012-0247-1

Geneen, L. J., Moore, R. A., Clarke, C., Martin, D., Colvin, L. A., & Smith, B. H. (2017). Physical activity and exercise for chronic pain in adults: An overview of Cochrane Reviews. *Cochrane Database Systematic Review*, *4*(4), Cd011279. https://doi.org/10.1002/14651858.CD011279.pub3

Gerdle, B., Molander, P., Stenberg, G., Stalnacke, B. M., & Enthoven, P. (2016). Weak outcome predictors of multimodal rehabilitation at one-year follow-up in patients with chronic pain—A practice based evidence study from two SQRP centres. *BMC Musculoskeletal Disorders*, *17*(1), 490. https://doi.org/10.1186/s12891-016-1346-7

Gerdle, B., Molander, P., Stenberg, G., Stålnacke, B.-M., & Enthoven, P. (2016). Weak outcome predictors of multimodal rehabilitation at one-year follow-up in patients with chronic pain—A practice based evidence study from two SQRP centres. *BMC Musculoskeletal Disorders*, *17*(1), 490. https://doi.org/10.1186/s12891-016-1346-7

Grahn, B., Ekdahl, C., & Borgquist, L. (2000). Motivation as a predictor of changes in quality of life and working ability in multidisciplinary rehabilitation. A two-year follow-up of a prospective controlled study in patients with prolonged musculoskeletal disorders. *Disability and Rehabilitation*, *22*(15), 639–654. https://doi.org/10.1080/096382800445443

Gustafsson, M., Ekholm, J., & Broman, L. (2002). Effects of a multi-professional rehabilitation programme for patients with fibromyalgia syndrome. *Journal of Rehabilitation Medicine*, *34*(3), 119–127. https://doi.org/10.1080/165019702753714147

Guzmán, J., Esmail, R., Karjalainen, K., Malmivaara, A., Irvin, E., & Bombardier, C. (2001). Multidisciplinary rehabilitation for chronic low back pain: Systematic review. *BMJ*, *322*(7301), 1511–1516. https://doi.org/10.1136/bmj.322.7301.1511

Hafenbrack, K., Heinrich, M., Müller, G., Marnitz, U., Mallwitz, J., & Klinger, R. (2013). Effects of interdisciplinary functional restoration treatment with cognitive behavior therapy in patients with chronic back pain: Healthcare research in the context of selective contracts. *Schmerz*, *27*(6), 566–576. https://doi.org/10.1007/s00482-013-1377-z (Effekte eines interdisziplinären Functional-restoration-Behandlungsprogramms mit kognitiv-behavioraler Therapie beim chronischen Rückenschmerz : Versorgungsforschung im Kontext von Selektivverträgen).

Haiduk, P., Benz, T., Lehmann, S., Gysi-Klaus, F., Aeschlimann, A., Michel, B. A., & Angst, F. (2017). Interdisciplinary rehabilitation after whiplash injury: An observational prospective 5 years outcome study. *Medicine (Baltimore)*, *96*(9), e6113. https://doi.org/10.1097/md.0000000000006113

Hållstam, A., Löfgren, M., Svensén, C., & Stålnacke, B. M. (2016). Patients with chronic pain: One-year follow-up of a multimodal rehabilitation programme at a pain clinic. *Scandinavian Journal of Pain*, *10*, 36–42. https://doi.org/10.1016/j.sjpain.2015.08.008

Hazard, R. G., Fenwick, J. W., Kalisch, S. M., Redmond, J., Reeves, V., Reid, S., & Frymoyer, J. W. (1989). Functional restoration with behavioral support. *Spine*, *14*(2), 157–161. https://doi.org/10.1097/00007632-198902000-00003

Higgins, J. P. T., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019)*. Cochrane.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hildebrandt, J., Pfingsten, M., Franz, C., Saur, P., & Seeger, D. (1996). Multidisciplinary treatment program for chronic low back pain, part 1. Overview. *Schmerz*, *10*(4), 190–203. https://doi.org/10.1007/s004829600018 (Das Göttinger Rücken Intensiv Programm (GRIP)—ein multimodales Behandlungsprogramm für Patienten mit chronischen Rückenschmerzen, Teil 1 Ergebnisse im Uberblick).

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch, P., Wyatt, J. C., Chan, A. W., & Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, *348*, g1687. https://doi.org/10.1136/bmj.g1687

Hozo, S. P., Djulbegovic, B., & Hozo, I. (2005). Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology*, *5*, 13. https://doi.org/10.1186/1471-2288-5-13

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, *11*(2), 193–206. https://doi.org/10.1037/1082-989x.11.2.193

Huffman, K. L., Mandell, D., Lehmann, J. K., Jimenez, X. F., & Lapin, B. R. (2019). Clinical and demographic predictors of interdisciplinary chronic pain rehabilitation program treatment response. *The Journal of Pain*, *20*(12), 1470–1485. https://doi.org/10.1016/j.jpain.2019.05.014

Ibrahim, M. E., Weber, K., Courvoisier, D. S., & Genevay, S. (2019). Recovering the capability to work among patients with chronic low Back pain after a four-week, multidisciplinary biopsychosocial rehabilitation program: 18-month follow-up study. *BMC Musculoskeletal Disorders*, *20*(1), 439. https://doi.org/10.1186/s12891-019-2831-6

Jeffery, M. M., Butler, M., Stark, A., & Kane, R. L. (2011). *Multidisciplinary pain programs for chronic noncancer pain. (Report No.11-EHC064-EF)*. Agency for Healthcare Research and Quality (US). Retrieved April 2, 2020, from https://www.ncbi.nlm.nih.gov/books/NBK82511/

Jensen, I. B., Dahlquist, C., Nygren, Å., Royen, E., Stenberg, M. (1997). Treatment for "helpless" women suffering from chronic spinal pain: A randomized controlled 18-month follow-up study. *Journal of Occupational Rehabilitation*, *7*(4), 225–238. https://doi.org/10.1023/B:JOOR.0000010858.70743.21

Kääpä, E. H., Frantsi, K., Sarna, S., & Malmivaara, A. (2006). Multidisciplinary group rehabilitation versus individual physiotherapy for chronic nonspecific low back pain. *Spine*, *31*(4), 371–376. https://doi.org/10.1097/01.brs.0000200104.90759.8c

Kaiser, U., Kopkow, C., Deckert, S., Neustadt, K., Jacobi, L., Cameron, P., De Angelis, V., Apfelbacher, C., Arnold, B., Birch, J., Bjarnegård, A., Christiansen, S., C de C Williams, A., Gossrau, G., Heinks, A., Hüppe, M., Kiers, H., Kleinert, U., Martelletti, P., ... Schmitt, J. (2018). Developing a core outcome domain set to assessing effectiveness of interdisciplinary multimodal pain therapy: The VAPAIN consensus statement on core outcome domains. *Pain*, *159*(4), 673–683. https://doi.org/10.1097/j.pain.0000000000001129

Kaiser, U., Treede, R. D., & Sabatowski, R. (2017). Multimodal pain therapy in chronic noncancer pain—Gold standard or need for further clarification? *Pain*, *158*(10), 1853–1859. https://doi.org/10.1097/j.pain.0000000000000902

Kamper, S. J., Apeldoorn, A. T., Chiarotto, A., Smeets, R. J., Ostelo, R. W., Guzman, J., & van Tulder, M. W. (2014). Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD000963.pub3

Koopman, F. S., Edelaar, M., Slikker, R., Reynders, K., van der Woude, L. H., & Hoozemans, M. J. (2004). Effectiveness of a multidisciplinary occupational training program for chronic low back pain: A prospective cohort study. *American Journal of Physical Medicine & Rehabilitation*, *83*(2), 94–103. https://doi.org/10.1097/01.Phm.0000107482.35803.11

Kösters, M. (2017). Every effect size has its place: A commentary on the avoidance of pre-post effect sizes. *Epidemiology and Psychiatric Sciences*, *26*(4), 369–370. https://doi.org/10.1017/s204579601700004x

Lemstra, M., & Olszynski, W. P. (2005). The effectiveness of multidisciplinary rehabilitation in the treatment of fibromyalgia: A randomized controlled trial. *Clinical Journal of Pain*, *21*(2), 166–174. https://doi.org/10.1097/00002508-200503000-00008

Letzel, J., Angst, F., & Weigl, M. B. (2019). Multidisciplinary biopsychosocial rehabilitation in chronic neck pain: A naturalistic prospective cohort study with intraindividual control of effects and 12-month follow-up. *European Journal of Physical and Rehabilitation Medicine*, *55*(5), 665–675. https://doi.org/10.23736/S1973-9087.18.05348-0

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage Publications.

Mangels, M., Schwarz, S., Worringen, U., Holme, M., & Rief, W. (2009). Evaluation of a behavioral-medical inpatient rehabilitation treatment including booster sessions: A randomized controlled study. *The Clinical Journal of Pain*, *25*(5), 356–364. https://doi.org/10.1097/AJP.0b013e3181925791

Martín, J., Torre, F., Padierna, A., Aguirre, U., González, N., García, S., Matellanes, B., & Quintana, J. M. (2012). Six-and 12-month follow-up of an interdisciplinary fibromyalgia treatment programme: Results of a randomised trial. *Clinical and Experimental Rheumatology*, *30*(6 Suppl 74), 103–111.

McAllister, M. J., McKenzie, K. E., Schultz, D. M., & Epshteyn, M. G. (2005). Effectiveness of a multidisciplinary chronic pain program for treatment of refractory patients with complicated chronic pain syndromes. *Pain Physician*, *8*(4), 369–373.

Meng, K., Seekatz, B., Roband, H., Worringen, U., Vogel, H., & Faller, H. (2011). Intermediate and long-term effects of a standardized back school for inpatient orthopedic rehabilitation on illness knowledge and self-management behaviors: A randomized controlled trial. *Clinical Journal of Pain*, *27*(3), 248–257. https://doi.org/10.1097/AJP.0b013e3181ffbfaf

Merrick, D., & Sjölund, B. H. (2009). Patients' pretreatment beliefs about recovery influence outcome of a pain rehabilitation program. *European Journal of Physical and Rehabilitation Medicine*, *45*(3), 391–401.

Merrick, D., Sundelin, G., & Stalnacke, B. M. (2012). One-year follow-up of two different rehabilitation strategies for patients with chronic pain. *Journal of Rehabilitation Medicine*, *44*(9), 764–773. https://doi.org/10.2340/16501977-1022

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

Monticone, M., Ambrosini, E., Rocca, B., Cazzaniga, D., Liquori, V., & Foti, C. (2016). Group-based task-oriented exercises aimed at managing kinesiophobia improved disability in chronic low back pain. *European Journal of Pain*, *20*(4), 541–551. https://doi.org/10.1002/ejp.756

Monticone, M., Ferrante, S., Rocca, B., Baiardi, P., Farra, F. D., & Foti, C. (2013). Effect of a long-lasting multidisciplinary program on disability and fear-avoidance behaviors in patients with chronic low back pain: Results of a randomized controlled

trial. *Clinical Journal of Pain*, *29*(11), 929–938. https://doi.org/10.1097/AJP.0b013e31827fef7e

Moola, S., Munn, Z., Tufanaru, C., Aromataris, E., Sears, K., Sfetcu, R., Currie, M., Qureshi, R., Mattis, P., Lisy, K., & Mu, P.-F. (2017). Chapter 7: Systematic reviews of etiology and risk. In E. Aromataris & Z. Munn (Eds.), *Joanna Briggs Institute reviewer's manual* (pp. 217–267). The Joanna Briggs Institute.

Morley, S. (2008). Relapse prevention: Still neglected after all these years. *Pain*, *134*(3), 239–240. https://doi.org/10.1016/j.pain.2007.12.004

Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology*, *53*(Pt 1), 17–29. https://doi.org/10.1348/000711000159150

Nagel, B., & Korb, J. J. (2009). Multimodale therapie. *Der Orthopäde*, *38*(10), 907–912. https://doi.org/10.1007/s00132-009-1482-y

Nicholas, M. K., Asghari, A., Sharpe, L., Brnabic, A., Wood, B. M., Overton, S., Tonkin, L., de Sousa, M., Finniss, D., Beeston, L., Sutherland, A., Corbett, M., & Brooker, C. (2014). Cognitive exposure versus avoidance in patients with chronic pain: Adherence matters. *European Journal of Pain*, *18*(3), 424–437. https://doi.org/10.1002/j.1532-2149.2013.00383.x

Olason, M. (2004). Outcome of an interdisciplinary pain management program in a rehabilitation clinic. *Work*, *22*(1), 9–15.

Opozda, M., Chur-Hansen, A., & Wittert, G. (2016). Changes in problematic and disordered eating after gastric bypass, adjustable gastric banding and vertical sleeve gastrectomy: A systematic review of pre-post studies. *Obesity Reviews*, *17*(8), 770–792. https://doi.org/10.1111/obr.12425

Oslund, S., Robinson, R. C., Clark, T. C., Garofalo, J. P., Behnk, P., Walker, B., Walker, K. E., Gatchel, R. J., Mahaney, M., & Noe, C. E. (2009). Long-term effectiveness of a comprehensive pain management program: Strengthening the case for interdisciplinary care. *Baylor University Medical Center Proceedings*, *22*(3), 211–214. https://doi.org/10.1080/08998280.2009.11928516

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Penney, L. S., & Haro, E. (2019). Qualitative evaluation of an interdisciplinary chronic pain intervention: Outcomes and barriers and facilitators to ongoing pain management. *Journal of Pain Research*, *12*, 865–878. https://doi.org/10.2147/jpr.S185652

Persson, E., Lexell, J., Eklund, M., & Rivano-Fischer, M. (2012). Positive effects of a musculoskeletal pain rehabilitation program regardless of pain duration or diagnosis. *Physical Medicine and Rehabilitation*, *4*(5), 355–366. https://doi.org/10.1016/j.pmrj.2011.11.007

Pietilä-Holmner, E., Enthoven, P., Gerdle, B., Molander, P., & Stålnacke, B. M. (2020). Long-term outcomes of multimodal rehabilitation in primary care for patients with chronic pain. *Journal of Rehabilitation Medicine*, *52*(2), jrm00023. https://doi.org/10.2340/16501977-2649

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Reck, T., Dumat, W., Krebs, J., & Ljutow, A. (2017). Outpatient multimodal pain therapy: Results of a 1-week intensive outpatient multimodal group program for patients with chronic unspecific back pain—Retrospective evaluation after 3 and 12 months. *Schmerz*, *31*(5), 508–515. https://doi.org/10.1007/s00482-017-0211-4 (Ambulante multimodale Schmerztherapie : Ergebnisse eines 1-wöchigen ambulanten intensiven multimodalen Gruppenprogramms für Patienten mit chronischen unspezifischen Rückenschmerzen—retrospektive Evaluation nach 3 und 12 Monaten).

Richardson, I. H., Richardson, P. H., Williams, A. C., Featherstone, J., & Harding, V. R. (1994). The effects of a cognitive-behavioral pain management programme on the quality of work and employment status of severely impaired chronic pain patients. *Disability and Rehabilitation*, *16*(1), 26–34. https://doi.org/10.3109/09638289409166433

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, *143*(2), 117–141. https://doi.org/10.1037/bul0000088

Roche-Leboucher, G., Petit-Lemanac'h, A., Bontoux, L., Dubus-Bausière, V., Parot-Shinkel, E., Fanello, S., Penneau-Fontbonne, D., Fouquet, N., Legrand, E., Roquelaure, Y., & Richard, I. (2011). Multidisciplinary intensive functional restoration versus outpatient active physiotherapy in chronic low back pain. *Spine*, *36*(26), 2235–2242. https://doi.org/10.1097/BRS.0b013e3182191e13

RStudio Team. (2020). *RStudio: Integrated development for R, PBC*. RStudio. http://www.rstudio.com/

Scascighini, L., Toma, V., Dober-Spielmann, S., & Sprott, H. (2008). Multidisciplinary treatment for chronic pain: A systematic review of interventions and outcomes. *Rheumatology*, *47*(5), 670–678. https://doi.org/10.1093/rheumatology/ken021

Semrau, J., Hentschke, C., Buchmann, J., Meng, K., Vogel, H., Faller, H., Bork, H., & Pfeifer, K. (2015). Long-term effects of interprofessional biopsychosocial rehabilitation for adults with chronic non-specific low back pain: A multicentre, quasi-experimental study. *PLoS One*, *10*(3), e0118609. https://doi.org/10.1371/journal.pone.0118609

Silvemark, A., Källmén, H., & Molander, C. (2014). Improved life satisfaction and pain reduction: Follow-up of a 5-week multidisciplinary long-term pain rehabilitation programme. *Upsala Journal of Medical Sciences*, *119*(3), 278–286. https://doi.org/10.3109/03009734.2014.908252

Smeets, R. J., Vlaeyen, J. W., Hidding, A., Kester, A. D., van der Heijden, G. J., & Knottnerus, J. A. (2008). Chronic low back pain: Physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain*, *134*(3), 263–276. https://doi.org/10.1016/j.pain.2007.04.021

Spinhoven, P., Ter Kuile, M., Kole-Snijders, A. M., Hutten Mansfeld, M., Den Ouden, D. J., & Vlaeyen, J. W. (2004). Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. *European Journal of Pain*, *8*(3), 211–219. https://doi.org/10.1016/j.ejpain.2003.08.003

Stein, K. F., & Miclescu, A. (2013). Effectiveness of multidisciplinary rehabilitation treatment for patients with chronic pain in a primary health care unit. *Scandinavian Journal of Pain*, *4*(4), 190–197. https://doi.org/10.1016/j.sjpain.2013.06.003

Steinmetz, A., Psczolla, M., Seidel, W., Niemier, K., Derlien, S., & Nisser, J. (2019). Effect of subgroup-specific multimodal

therapy on chronic spinal back pain and function—A prospective inpatient multicentre clinical trial in Germany. *Medicine*, *98*(1), e13825. https://doi.org/10.1097/md.0000000000013825

Strobel, E. S., Wild, J., & Müller, W. (1998). Interdisciplinary group therapy for fibromyalgia. *Zeitschrift fur Rheumatologie*, *57*(2), 89–94. https://doi.org/10.1007/s003930050065 (Interdisziplinäre Gruppentherapie für die Fibromyalgie).

Tardif, H., Arnold, C., Hayes, C., & Eagar, K. (2017). Establishment of the Australasian electronic persistent pain outcomes collaboration. *Pain Medicine*, *18*(6), 1007–1018. https://doi.org/10.1093/pm/pnw201

Tavafian, S. S., Jamshidi, A. R., & Mohammad, K. (2011). Treatment of chronic low back pain: A randomized clinical trial comparing multidisciplinary group-based rehabilitation program and oral drug treatment with oral drug treatment alone. *The Clinical Journal of Pain*, *27*(9), 811–818. https://doi.org/10.1097/AJP.0b013e31821e7930

Thieme, K., Gromnica-Ihle, E., & Flor, H. (2003). Operant behavioral treatment of fibromyalgia: A controlled study. *Arthritis & Rheumatism*, *49*(3), 314–320. https://doi.org/10.1002/art.11124

Thomas, J., Noel-Storr, A., Marshall, I., Wallace, B., McDonald, S., Mavergames, C., Glasziou, P., Shemilt, I., Synnot, A., Turner, T., Elliott, J., Agoritsas, T., Hilton, J., Perron, C., Akl, E., Hodder, R., Pestridge, C., Albrecht, L., Horsley, T., ... Pearson, L. (2017). Living systematic reviews: 2. Combining human and machine effort. *Journal of Clinical Epidemiology*, *91*, 31–37. https://doi.org/10.1016/j.jclinepi.2017.08.011

Treede, R. D., Rief, W., Barke, A., Aziz, Q., Bennett, M. I., Benoliel, R., Cohen, M., Evers, S., Finnerup, N. B., First, M. B., Giamberardino, M. A., Kaasa, S., Kosek, E., Lavand'homme, P., Nicholas, M., Perrot, S., Scholz, J., Schug, S., Smith, B. H., ... Wang, S. J. (2015). A classification of chronic pain for ICD-11. *Pain*, *156*(6), 1003–1007. https://doi.org/10.1097/j.pain.0000000000000160

Treede, R. D., Rief, W., Barke, A., Aziz, Q., Bennett, M. I., Benoliel, R., Cohen, M., Evers, S., Finnerup, N. B., First, M. B., Giamberardino, M. A., Kaasa, S., Korwisi, B., Kosek, E., Lavand'homme, P., Nicholas, M., Perrot, S., Scholz, J., Schug, S., ... Wang, S. J. (2019). Chronic pain as a symptom or a disease: The IASP classification of chronic pain for the International Classification of Diseases (ICD-11). *Pain*, *160*(1), 19–27.

Turk, D. C. (2003). Cognitive-behavioral approach to the treatment of chronic pain patients. *Regional Anesthesia and Pain Medicine*, *28*(6), 573–579. https://doi.org/10.1016/s1098-7339(03)00392-4

Turk, D. C., Dworkin, R. H., Allen, R. R., Bellamy, N., Brandenburg, N., Carr, D. B., Cleeland, C., Dionne, R., Farrar, J. T., Galer, B. S., Hewitt, D. J., Jadad, A. R., Katz, N. P., Kramer, L. D., Manning, D. C., McCormick, C. G., McDermott, M. P., McGrath, P., Quessy, S., ... Witter, J. (2003). Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain*, *106*(3), 337–345. https://doi.org/10.1016/j.pain.2003.08.001

Turk, D. C., & Rudy, T. E. (1991). Neglected topics in the treatment of chronic pain patients—relapse, noncompliance, and adherence enhancement. *Pain*, *44*(1), 5–28. https://doi.org/10.1016/0304-3959(91)90142-k

Van De Velde, D., De Zutter, F., Satink, T., Costa, U., Janquart, S., Senn, D., & De Vriendt, P. (2019). Delineating the concept of self-management in chronic conditions: A concept analysis. *British Medical Journal Open*, *9*(7). https://doi.org/10.1136/bmjopen-2018-027775

Van der Maas, L. C., Köke, A., Pont, M., Bosscher, R. J., Twisk, J. W., Janssen, T. W., & Peters, M. L. (2015). Improving the multidisciplinary treatment of chronic pain by stimulating body awareness: A cluster-randomized trial. *Clinical Journal of Pain*, *31*(7), 660–669. https://doi.org/10.1097/ajp.0000000000000138

van Hooff, M. L., van der Merwe, J. D., O'Dowd, J., Pavlov, P. W., Spruit, M., de Kleuver, M., & van Limbeek, J. (2010). Daily functioning and self-management in patients with chronic low back pain after an intensive cognitive behavioral programme for pain management. *European Spine Journal*, *19*(9), 1517–1526. https://doi.org/10.1007/s00586-010-1435-5

van Wilgen, C. P., Dijkstra, P. U., Versteegen, G. J., Fleuren, M. J., Stewart, R., & van Wijhe, M. (2009). Chronic pain and severe disuse syndrome: Long-term outcome of an inpatient multidisciplinary cognitive behavioural programme. *Journal of Rehabilitation Medicine*, *41*(3), 122–128. https://doi.org/10.2340/16501977-0292

Vendrig, A. A., Hoofs, M. H., van Akkerveeken, P. F., & Lamberts-Hopkes, K. J. (2000). Multidisplinary approach to chronic back pain: Postrehabilitation resumption of work the same 3-4 years later as after 6 months. *The Nederlands Tijdschrift Voor Geneeskunde*, *144*(46), 2207–2209 (Multidisciplinaire aanpak van chronische rugpijn: Werkhervatting na 3-4 jaar hetzelfde als na 6 maanden).

Verkerk, K., Luijsterburg, P. A., Ronchetti, I., Miedema, H. S., Pool-Goudzwaard, A., van Wingerden, J. P., & Koes, B. W. (2011). Course and prognosis of recovery for chronic non-specific low back pain: Design, therapy program and baseline data of a prospective cohort study. *BMC Musculoskeletal Disorders*, *12*, 252. https://doi.org/10.1186/1471-2474-12-252

Viechtbauer. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vlaeyen, J. W., Crombez, G., & Linton, S. J. (2016). The fear-avoidance model of pain. *Pain*, *157*(8), 1588–1589. https://doi.org/10.1097/j.pain.0000000000000574

Volker, G., van Vree, F., Wolterbeek, R., van Gestel, M., Smeets, R., Köke, A., & Vlieland, T. V. (2017). Long-term outcomes of multidisciplinary rehabilitation for chronic musculoskeletal pain. *Musculoskeletal Care*, *15*(1), 59–68. https://doi.org/10.1002/msc.1141

Vowles, K. E., McCracken, L. M., & O'Brien, J. Z. (2011). Acceptance and values-based action in chronic pain: A three-year follow-up analysis of treatment effectiveness and process. *Behaviour Research and Therapy*, *49*(11), 748–755. https://doi.org/10.1016/j.brat.2011.08.002

Wagner, F., Janzen, B., Tkachuk, G., Laverty, W., & Woods, M. J. (2011). An evaluation of the long-term treatment outcomes of an interdisciplinary chronic pain centre program. *Journal of Musculoskeletal Pain*, *19*(3), 134–143. https://doi.org/10.3109/10582452.2011.582985

Waterschoot, F. P., Dijkstra, P. U., Hollak, N., de Vries, H. J., Geertzen, J. H., & Reneman, M. F. (2014). Dose or content? Effectiveness of pain rehabilitation programs for patients with chronic low back pain: A systematic review. *Pain*, *155*(1), 179–189. https://doi.org/10.1016/j.pain.2013.10.006

Westman, A., Linton, S. J., Theorell, T., Ohrvik, J., Wahlén, P., & Leppert, J. (2006). Quality of life and maintenance of

improvements after early multimodal rehabilitation: A 5-year follow-up. *Disability and Rehabilitation*, *28*(7), 437–446. https://doi.org/10.1080/09638280500192694

Williams, A. C., Eccleston, C., & Morley, S. (2012). Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database Systematic Review*, *11*(11), Cd007407. https://doi.org/10.1002/14651858.CD007407.pub3

Williams, A. C., Richardson, P. H., Nicholas, M. K., Pither, C. E., Harding, V. R., Ridout, K. L., Ralphs, J. A., Richardson, I. H., Justins, D. M., & Chamberlain, J. H. (1996). Inpatient vs. outpatient pain management: Results of a randomised controlled trial. *Pain*, *66*, 13–22. https://doi.org/10.1016/0304-3959(96)02996-X

Williamson, P. R., Altman, D. G., Bagley, H., Barnes, K. L., Blazeby, J. M., Brookes, S. T., Clarke, M., Gargon, E., Gorst, S., Harman, N., Kirkham, J. J., McNair, A., Prinsen, C. A. C., Schmitt, J., Terwee, C. B., & Young, B. (2017). The COMET handbook: Version 1.0. *Trials*, *18*(Suppl 3), 280. https://doi.org/10.1186/s13063-017-1978-4

Wilson, G. T. (2010). Eating disorders, obesity and addiction. *European Eating Disorders Review*, *18*(5), 341–351. https://doi.org/10.1002/erv.1048

Wood, W., & Neal, D. T. (2016). Healthy through habit: Interventions for initiating & maintaining health behavior change. *Behavioral Science & Policy*, *2*(1), 71–83.

Zhuk, A., Schiltenwolf, M., & Neubauer, E. (2018). Long-term efficacy of multimodal pain therapy for chronic back pain. *Nervenarzt*, *89*(5), 546–551. https://doi.org/10.1007/s00115-017-0391-2 (Langfristige Wirksamkeit einer multimodalen Schmerztherapie bei chronischen Rückenschmerzen).

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Elbers, S., Wittink, H., Konings, S., Kaiser, U., Kleijnen, J., Pool, J., Köke, A., & Smeets, R. (2022). Longitudinal outcome evaluations of Interdisciplinary Multimodal Pain Treatment programmes for patients with chronic primary musculoskeletal pain: A systematic review and meta-analysis. *European Journal of Pain*, 26, 310–335. https://doi.org/10.1002/ejp.1875