Reported Outcomes in Interdisciplinary Pain Treatment: An Overview of Systematic Reviews and Meta-Analyses of Randomised Controlled Trials

Huan-Ji Dong, Björn Gerdle, Elena Dragioti

Pain and Rehabilitation Centre, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

Correspondence: Huan-Ji Dong, Pain and Rehabilitation Centre, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden, Email huanji.dong@liu.se

Background: There is considerable diversity of outcome selections and methodologies for handling the multiple outcomes across all systematic reviews (SRs) of Interdisciplinary Pain Treatment (IPT) due to the complexity. This diversity presents difficulties for healthcare decision makers. Better recommendations about how to select outcomes in SRs (with or without meta-analysis) are needed to explicitly demonstrate the effectiveness of IPT.

Objective: This overview systematically collates the reported outcomes and measurements of IPT across published SRs and identifies the methodological characteristics. Additionally, we provide some suggestions on framing the selection of outcomes and on conducting SRs of IPT.

Methods: Three electronic databases (PubMed, Cochrane Library, and Epistemonikos) and the PROSPERO registry for ongoing SR were supplemented with hand-searching ending on 30 September 2021.

Results: We included 18 SRs with data on 49,007 people from 356 primary randomised controlled trials (RCTs); eight were followed by meta-analysis and ten used narrative syntheses of data. For all the SRs, pain was the most common reported outcome (72%), followed by disability/function status (61%) and working status (61%). Psychological well-being and quality of life were also reported in half of the included SR (50%). The core outcome domains according to VAPAIN, IMMPACT, and PROMIS were seldom met. The methodological quality varied from critically low to moderate according to AMSTAR2. The AMSTAR2 rating was negatively correlated to the number of outcome domains in PROMIS, and VAPAIN was positively correlated with IMMPACT and PROMIS, indicating the intercorrelations between the reported outcomes.

Conclusion: This systematic overview showed wide-ranging disparity in reported outcomes and applied outcome domains in SRs evaluating IPT interventions for chronic pain conditions. The intercorrelations between the reported outcomes should be appropriately handled in future research. Some approaches are discussed as well.

Keywords: interdisciplinary pain treatment, interdisciplinary pain rehabilitation program, multimodal rehabilitation, multidisciplinary rehabilitation, biopsychosocial pain rehabilitation, pain management program, chronic pain, outcome domains, intercorrelation

Introduction
Description of the Patient Populations
According to the International Association for the Study of Pain (IASP), chronic pain is pain that persists or recurs for more than 3 months. Significant emotional distress and/or functional disability are well-known characteristics of chronic pain. The prevalence of chronic non-cancer pain is estimated to be between 8.7% and 64.4% across diverse populations around the world, depending on the assessments method used. In Europe, about 19% adults suffer from chronic pain conditions with moderate-severe pain intensity. In the US, about 8% of adults have chronic pain that cause limitations in normal life activities such as work, social, recreational, and self-care activities. When assessing the patient with chronic pain, it is important to determine the activated pain mechanisms, ie, if the pain is nociceptive (injury to non-neural tissues producing noxious stimulus), neuropathic (injury to the somatosensory nervous system), nociceplastic (altered nociception...
Despite no clear evidence of tissue damage or combinations of these.\textsuperscript{5} Such characterisations are important steps towards mechanism-based management and treatments of chronic pain conditions.

**Description of Interdisciplinary Pain Treatment (IPT)**

Interdisciplinary pain treatment (IPT), based on the IASP, is a multimodal treatment provided by a multidisciplinary team collaborating in assessment and treatment using a shared bio-psycho-social (BPS) model and goals.\textsuperscript{6} This model suggests that chronic pain is based on a foundation of neurobiology (partially unknown) influenced by and interacting with biological, psychological, and social/contextual factors.\textsuperscript{7–11} The BPS framework is applied in modern clinical management of pain to better evaluate pain and its negative consequences for patients,\textsuperscript{12,13} including physical functioning,\textsuperscript{11} psychological well-being,\textsuperscript{16} and quality of life.\textsuperscript{14} Apart from medical treatment, IPT usually contains two or more well-synchronised treatment components such as physical, occupational, psychological, social, and educational components over a specified period (several weeks). Unlike pharmacological treatments or unimodal interventions, IPT is a complex intervention targeting the whole patient, including behaviours.\textsuperscript{15}

Besides IPT, other common names used in the literature are Interdisciplinary pain rehabilitation program (IPRP), multimodal rehabilitation, multidisciplinary rehabilitation, biopsychosocial pain rehabilitation, and pain management program.

**How IPT Works**

IPT is generally offered when unimodal interventions (eg, pharmacological treatment, surgery, and rehabilitation with a unimodal approach) have not been associated with important improvements or when the patient with chronic pain has a complex condition, for example, with respect to comorbidities (eg, depressive and anxiety symptoms) and/or prolonged sick leave and difficulties with return to work. The BPS framework captures the total complexity of the pain condition and provides clues for reasonable interventions, including IPT. Hence, IPT targets the complexity in several components simultaneously and is well-coordinated (pain intensity, depressive symptoms, perceived health, and return to work) over time. For years, several systematic reviews (SRs) have reported that IPT is more effective compared with single-treatment or usual care.\textsuperscript{16–18} However, how this is achieved in detail and the importance of duration or other dosage aspects are not known.\textsuperscript{19}

**Why It is Important to Do This Review**

True measures of healthcare quality are the outcomes that matter to patients.\textsuperscript{20} By using patient-reported outcomes (PROMs), we can measure the quality of IPT and identify possible ways to improve IPT.\textsuperscript{21} Pain is a personal experience and patients with chronic pain conditions often describe wide consequences: significant pain intensity, psychological distress, insomnia, reduced work ability and sick-leave, ill-health, worse physical functioning and low quality of life.\textsuperscript{22–27} The large variation of the study procedures and published outcomes makes comparing RCTs difficult.\textsuperscript{28} Without standardised outcome measures, interpretation and implementation of evidence are challenging. To improve the choice of outcome measures to assess the effectiveness of IPT, several core outcomes sets (COSs) have been developed. In this paper, we used three relatively established guidelines for reporting outcomes: 1) Validation and application of a core set of patient-relevant outcome domains (V APAIN);\textsuperscript{29} 2) Initiative on methods, measurements, and pain assessments in clinical trials (IMMPACT);\textsuperscript{30} and 3) Patient-reported outcomes measurement information system (PROMIS).\textsuperscript{31}

IPT is a complex intervention. The published systematic reviews (SRs) that explicitly select IPT outcomes have not been well studied and documented.\textsuperscript{32} Therefore, the evaluation of complex interventions such as IPT are ambiguous and different definitions of a positive outcome of an IPT trial have been presented.\textsuperscript{33} On the other hand, meta-research studies (ie, research on research) have increased rapidly to evaluate and improve research methods and reporting practices.\textsuperscript{34} To date, only one SR of randomised controlled trials (RCTs) has been used to investigate reported outcomes of an IPT,\textsuperscript{28} and no study has determined which reported outcomes are used in SRs of IPT. To fill this knowledge gap, we aim to provide a meticulous overview of the IPT outcomes reported in SRs with or without meta-analysis to evaluate the quality of IPT reported outcomes. In addition, we provide an overview of methodological aspects related to SRs of IPT to initiate a discussion of how to develop new methods of conducting and reporting SRs with complex interventions.
Methods

Study Design

We conducted a systematic overview of reviews\textsuperscript{35} following a standardised methodological approach in line with a published protocol (International Registered Report Identifier: PRR1-10.2196/17795).\textsuperscript{32} We also reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statement, including the PRISMA 2020 for Abstracts checklist\textsuperscript{36} (see PRISMA Tables S1 and S2 in the Supplementary materials).

Search Strategy

We undertook a comprehensive literature search using PubMed, Cochrane Database of Systematic Reviews, and Epistemonikos from the earliest record through 30 September 2021. A search strategy was created using a combination of the following terms adjusted per database if needed: chronic pain, neuropathic pain, chronic persistent pain, pain rehabilitation, pain therapy, multidisciplinary, interdisciplinary, multimodal, multidisciplinary biopsychosocial rehabilitation, combined modality therapy, and patient care team. Hence, with respect to selection of pain conditions, we included nociceptive, neuropathic and nociplastic chronic pain conditions. We also used a sensitive filter created to retrieve only SRs citations in PubMed.\textsuperscript{37} The full search strategy is outlined in Box S1 in the Supplementary material. The search strategy was limited by study design, but no restrictions were made on setting or context of the included articles or publication date. We also screened the reference lists of relevant records for additional articles and searched the International Prospective Register of Systematic Reviews (PROSPERO) for relevant ongoing reviews (Table S3 in the Supplementary material). Two independent reviewers (ED and H-JD) searched the databases, evaluated the titles and abstracts of each identified record, examined the full text eligibility, and assessed whether a potentially eligible article fulfilled our inclusion criteria. Any disagreement was solved by discussion between the two reviewers or, if required, a third reviewer (BG) was consulted until agreement was reached.

Study Selection

We included SRs with or without meta-analysis assessing the effectiveness of IPT strategies for chronic pain. SRs that assessed IPT as a part of other interventions were included if separate results and outcomes for IPT were given. We also included other types of evidence synthesis that have been published and fulfilled our following inclusion criteria – ie, network meta-analyses or living SRs, individual patient data meta-analysis, and health agency reviews.

To be eligible, SRs needed to accomplish the following: 1) present a definition of IPT throughout the full text along with description of components involved in the intervention; 2) include adult male or female participants or both (at least 75% of the participants must be \( \geq 18 \) years of age) with chronic/persistent nociceptive, neuropathic and/or nociplastic pain condition (ie, pain lasting at least 3 months) as described in our protocol;\textsuperscript{32} 3) include only RCTs published in English or Swedish; and 4) compare IPT with treatment as usual, wait list, no treatment, or an alternative active intervention including comparisons between two different types of IPT.

Although there is some overlap, SRs that assessed “back school interventions” or work-hardening programmes were excluded because they are not considered identical treatment to IPT.\textsuperscript{38} We further excluded SRs if they had any of the following characteristics: 1) included patients with a diagnosis of chronic pain due to cancer, infection, inflammatory arthropathy, osteoporosis, fracture, pregnancy, rheumatoid arthritis, or other rheumatic pain;\textsuperscript{32} 2) included patients with subacute or acute pain (ie, pain less than 3 months); 3) included study designs other than RCTs (eg, clinical trials without randomization, single study designs, and cohort studies); 4) included secondary research studies as the unit of analysis (ie, overview of reviews, umbrella reviews, scoping reviews, summary of systematic reviews, or mixed methods review); and 5) published in languages other than English or Swedish. In cases of duplicate publications (ie, the same SR from the same research group in a different journal) (see\textsuperscript{38,39}), we kept the most comprehensive and detailed version of the article. However, updated versions of the same SRs were considered for inclusion (see\textsuperscript{38,40}).
Data Extraction
Endnote software (X9 version) was used to collect the results of the searches and to remove duplicates. Two independent reviewers (ED and H-JD) performed data extraction using pre-defined excel forms. The list of variables of interest included the following: PMID/DOI of the included SR, first author, publication year, chronic pain conditions, control/comparison arms, number of RCTs of IPTs included in the SR, outcomes investigated (primary and secondary if such categorization exists), and total number of participants. Furthermore, we abstracted data in terms of the duration of the treatment (weeks and hours), treatment components, setting, and follow-up length. Finally, we recorded any method, strategy, considerations, or discussion regarding how the authors chose which outcomes to study and which methods to use to evaluate the evidence (eg, the GRADE approach) as well as methodological aspects such as reporting of funding, methods for risk of bias appraisal, synthesis of results, and details of meta-analytic processes if meta-analysis were conducted.

Any discrepancy was solved by discussion between the two reviewers. Data extraction on characteristics of included studies and reported outcomes of ITP was double checked by a third reviewer (BG).

Quality Assessment of Included Reviews
The quality of the included SRs was rated by means of the AMSTAR 2 tool (A Measurement Tool to Assess Systematic Reviews version 2)\(^1\) and by two independent reviewers (ED and H-JD). Any discrepancy between two reviewers for individual AMSTAR-2 questions was discussed until agreement was reached. AMSTAR-2 comprises 16 items and four options are provided to answer the questions of items: “Yes”, “Partial Yes”, “No” or “No meta-analysis conducted”. Seven critical items involve protocol content and registration, comprehensive literature search, providing reasons for excluding studies, assessment of RoB for individual studies, appropriate synthesis of results (meta-analytic methods), consideration of RoB in results and impact of publication bias. AMSTAR 2 categorises the quality of systematic reviews into four domains – high quality, moderate quality, low quality, or critically low quality – based on 16 items and sets a very high threshold for a high overall rating.\(^2\)

Data Synthesis
We performed descriptive statistics to present our results using median and interquartile range for quantitative variables and absolute and relative frequencies for qualitative variables. We also used a narrative synthesis approach employing tabular presentations stratified by the AMSTAR2 assessment per SR. For conceptual classification, outcome domains were organised using three established guidelines of reporting outcomes: 1) VAPAIN statement for IPTs including eight core outcome domains;\(^3\) 2) IMMPACT statement including six core outcome domains;\(^4\) and 3) PROMIS recommendations including three core outcome domains.\(^5\) Each domain per statement was scored as positive (Y) or negative (N). The total score was computed by counting the number of outcome domains scored as positive. For AMSTAR2, we rated critically low quality as one, low quality as two, moderate quality as three, and high quality as four. Lastly, we used the spearman rho correlation coefficient (0.00–0.19: very weak; 0.20–0.39: weak; 0.40–0.59: moderate; 0.60–0.79: strong; ≥0.80: very strong) to determine the relationship between the AMSTAR2 rating and the other variables such as number of outcome domains in VAPAIN, IMMPACT, or PROMIS, total number of outcomes, total sample size, number of included studies, and year of publication. All analyses were conducted in STATA, version 17.0 (StatCorp LLC) and R software, version 4.1.2.

Differences Between Protocol and Review
We made small changes in terms of the inclusion criteria between this version and the published protocol.\(^6\) This involved clarifying that other types of SRs as described in study selection were considered for inclusion and excluding the criterion of the detailed description of professionals involved in the intervention as only a trivial number of the included SRs provided this information in detail. We also updated our search up to 30 September 2021, including PROSPERO database.
Results
Search Results
Figure 1 illustrates the PRISMA flow of the selection process. After deduplication, the initial electronic database search identified 3253 potentially eligible systematic reviews. After screening records by title and abstract, we considered 145 potentially eligible reviews for inclusion and retrieved full-text articles. Grounds for exclusion after full-text assessment are illustrated in Table S3 in the Supplementary material. We also identified six ongoing reviews on PROSPERO; three were completed and published,43–45 but only one45 met our inclusion criteria and therefore was included in the final set (see Table S2 in the Supplementary material for details). Finally, 18 published articles met the inclusion criteria and were included in this review.16–18,38,40,45–57

Characteristics of Included IPT Systematic Reviews
The 18 selected SRs included a total of 356 primary RCTs and a total of 49006 participants. Eight of the included SRs were quantitative syntheses followed by meta-analysis and ten used narrative syntheses of data. Most of the SRs focused largely or completely on chronic low back pain and three on fibromyalgia. The remaining focused on chronic non-cancer

Figure 1 PRISMA flow chart.
Notes: Adapted from: Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71. Creative Commons Attribution (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/legalcode).
pain in general. No SR eligible for our overview focused specifically on neuropathic pain. The proportion of female participants ranged between 21% and 100%. The median number of the total sample size per SR was 2212 (inter-quartile range [IQR] = 1334 to 4925, range 600 to 6858) and the median number of included primary RCTs was 15.5 (IQR = 9 to 27, range 2 to 46). A comprehensive description of each SR is provided in Tables 1 and 2. The methodological quality varied from critically low to moderate according to AMSTAR2 (see Table S4 in the Supplementary material for details).

The duration of the IPT treatment of the RCTs evaluated varied considerably across SRs (ranged from 1 week to 24 weeks) and only two SRs clearly reported the hours of treatment. All but three SRs included RCTs that compared an IPT with both active and passive comparison arms (eg, treatment as usual (TAU), waiting list control (WLC), no intervention, No-IPT interventions, and different types of IPTs). Details of treatment components and professionals involved were rarely reported in SRs. Only six gave details about population settings for the primary RCTs (eg, outpatients or inpatients, and primary care or specialist care).

Characteristics of Reported Outcomes in IPT Systematic Reviews

In all included SRs, the reported outcomes were assessed by validated self-report instruments with few exceptions for work status outcomes, which used data from labour markets, social security systems, or unemployment rates. Only four SRs (three SRs with meta-analysis and one without) clearly stated primary and secondary outcomes. In one SR, the authors specified only measures collected at long-term follow-up (12 months or more) were considered the primary outcomes (Tables 1 and 2). The median number of reported outcomes per SR was 5 (IQR = 3 to 7, range 1 to 14). The most common reported outcomes were pain (72%), disability or functional status (61%), work status (61%), psychological or emotional strain and cognitive function (50%), and quality of life (50%). However, only seven SRs provided a specific strategy for selecting outcomes (39%). Notably, only two SRs (11%) assessed outcome domains using established guidelines (OMERACT-10 and IMMACT) of reporting outcomes (Tables 1 and 2).

Reported Outcomes Domains Measured by VAPAIN, IMMPACT, and PROMIS Guidelines in IPT Systematic Reviews

According to VAPAIN, no SR assessed a combination of all eight core outcome domains (Table 3). The median number of reported outcome domains per SR was 3.5 (IQR = 1 to 5, range 1 to 7). Physical activity (14/18 SRs, 78%), pain intensity (12/18 SRs, 67%), productivity (11/18 SRs, 61%), emotional well-being (10/18 SRs, 55%), and health-related quality of life (9/18 SRs, 50%) were most frequently reported. The less commonly reported outcome domains were satisfaction with social roles and the patient’s activities and perceptions of treatment goal achievement (both 4/18 SRs, 22%). Only two SRs assessed a combination of all six core outcome domains according to IMMPACT (Table 4). The median number of reported outcome domains per SR was 2.5 (IQR = 1 to 3, range 1 to 6). Physical functioning (15/18 SRs, 83%), pain (12/18 SRs, 67%), and emotional functioning (11/18 SRs, 61%) were most frequently reported. The less commonly reported outcome domains were symptoms and adverse events (5/18 SRs, 28%) and participant disposition (2/18 SRs, 11%).

According to PROMIS, six SRs assessed a combination of all three core outcome domains (Table 4). The median number of reported outcome domains per SR was 2 (IQR = 1 to 3, range 1 to 3). Physical health was reported in 15/18 SRs (83%), social health in 11/18 SRs (61%), and mental health in 10/18 SRs (55%).

Univariate Spearman rho analysis (Figure 2) showed a negative correlation between AMSTAR2 rating and PROMIS score (weak) as well as AMSTAR2 rating and total number of outcomes (moderate). VAPAIN score was positively correlated with IMMPACT score and PROMIS score (very strong). The scores for IMMPACT, VAPAIN, and PROMIS had negative correlations with year of publication (strong and moderate). Total sample size was positively correlated with number of included studies (very strong).

Methodological Characteristics of Included IPT Systematic Reviews

None of the SRs assessed IPT as a complex intervention. However, 11 SRs reported the quality of the evidence with an appropriate guidance and comprehensive methods. Likewise, two other SRs used simple methods (ie, “the higher effectiveness had to be demonstrated in at least two out of the five primary outcomes, or at least
Table 1 Characteristics of Systematic Reviews with Meta-Analysis Stratified by AMSTAR2 Evaluation

<table>
<thead>
<tr>
<th>Author, Year (Ref)/Objectives</th>
<th>Pain Condition Treated (Definition)/Setting</th>
<th>IPT (Definition)</th>
<th>No. of Included Studies (Total Sample Size)</th>
<th>Mean Age or Age Range, Years (% Female)</th>
<th>Outcomes, (No. of Reported Outcomes)</th>
<th>Distinction Between Primary and Secondary</th>
<th>Specific Strategy Used to Select Outcomes</th>
<th>Risk of Bias Assessment in Included Studies</th>
<th>Evaluation of the Evidence</th>
<th>AMSTAR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casey, 2020†† the effectiveness of IPT versus active physical interventions</td>
<td>Chronic non-cancer pain (pain persisting for 12 wks or more)/NR</td>
<td>A coordinated approach to pain management, usually delivered to groups by a team of health care professionals from disciplines such as pain medicine, nursing, psychology, physiotherapy, and occupational therapy</td>
<td>27 (4424)</td>
<td>40 to 61 (NR)</td>
<td>Pain intensity, disability/ (2)</td>
<td>No</td>
<td>If &gt;1 measure of pain intensity or disability was used in a study, then the one considered to be the primary outcome</td>
<td>Serious high risk of bias in the majority of included studies</td>
<td>GRADE</td>
<td>Moderate</td>
</tr>
<tr>
<td>Martinez-Calderon, 2020†† which interventions enhance pain self-efficacy</td>
<td>Chronic musculoskeletal pain, according to the multidimensional diagnostic criteria for chronic pain by Dworkin et al, 2016†/NR</td>
<td>Interventions involving a combination of therapies (ie, exercise plus psychological therapy or self-management strategies plus exercise)</td>
<td>32 (5425)</td>
<td>NR (NR)</td>
<td>Pain self-efficacy/(1)</td>
<td>No</td>
<td>No</td>
<td>Serious high risk of bias in the majority of included studies</td>
<td>GRADE</td>
<td>Moderate</td>
</tr>
<tr>
<td>Papadopoulou, 2016†† the effectiveness of pharmacological and non-pharmacological interventions</td>
<td>Fibromyalgia syndrome (based on ACR 1990 criteria)/NR</td>
<td>A set of therapeutic strategies to manage pain</td>
<td>8 (1027)</td>
<td>NR (NR)</td>
<td>Pain intensity, sleep disturbance, physical function, fatigue, depression, anxiety, and dyscognition/ (7)</td>
<td>No</td>
<td>OMERACT-10</td>
<td>Serious high risk of bias in the majority of included studies</td>
<td>No</td>
<td>Moderate</td>
</tr>
<tr>
<td>Kamper, 2014†† the effectiveness of IPT versus usual care and with physical treatments</td>
<td>Chronic low back pain (pain persisting for 12 wks)/ Rehabilitation units</td>
<td>A physical component (eg, an exercise program) and at least one other element from the biopsychosocial model that is psychological or social and occupational.</td>
<td>41 (6858)</td>
<td>40 to 45 (21 to 100)</td>
<td>Pain, back-specific disability or functional status, and work status (return to work, sick leave), generic health or quality of life, healthcare service utilisation, global improvement, psychological and cognitive function (depression, anxiety, fear avoidance, coping strategies), adverse events/ (12)</td>
<td>Yes</td>
<td>Measures collected at long-term follow-up were considered primary outcomes</td>
<td>Thirteen studies were assessed as low risk of bias</td>
<td>GRADE</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Author, Year (Ref)/Objectives</th>
<th>Pain Condition Treated (Definition)/Setting</th>
<th>IPT (Definition)</th>
<th>No. of Included Studies (Total Sample Size)</th>
<th>Mean Age or Age Range, Years (% Female)</th>
<th>Outcomes, (No. of Reported Outcomes)</th>
<th>Distinction Between Primary and Secondary</th>
<th>Specific Strategy Used to Select Outcomes</th>
<th>Risk of Bias Assessment in Included Studies</th>
<th>Evaluation of the Evidence</th>
<th>AMSTAR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>van Middelkoop, 2011&lt;sup&gt;49&lt;/sup&gt; the effectiveness of physical and rehabilitation interventions</td>
<td>Chronic low back pain (defined as pain persisting for 12 wks)/NR</td>
<td>One physical dimension and one of the other dimensions (psychological or social or occupational)</td>
<td>6 (1229)</td>
<td>38 to 43 (37 to 100)</td>
<td>Pain, physical functional status, perceived recovery (eg, overall improvement), return to work (eg, return to work status, sick leave days), and side effects/ (5)</td>
<td>Yes</td>
<td>No</td>
<td>Three studies were assessed as low risk of bias</td>
<td>GRADE</td>
<td>Low</td>
</tr>
<tr>
<td>Häuser, 2009&lt;sup&gt;49&lt;/sup&gt; the efficacy of IPT</td>
<td>Fibromyalgia syndrome (based on ACR 1990 criteria)/Outpatients</td>
<td>The therapy treatment had to include at least two non-pharmacologic therapies (at least one educational or other psychological therapy and at least one exercise therapy)</td>
<td>9 (1119)</td>
<td>44 to 50 (96)</td>
<td>Pain, fatigue, sleep disturbances, depressive symptoms, and quality of life/ (5)</td>
<td>No</td>
<td>No</td>
<td>Six studies were of moderate quality and three studies were of low quality</td>
<td>No</td>
<td>Low</td>
</tr>
<tr>
<td>Gianola, 2018&lt;sup&gt;49&lt;/sup&gt; to compare via calculation of minimal important difference units, outcomes after IPT versus usual care</td>
<td>Chronic low back pain (defined as pain persisting for 3 mo or more)/NR</td>
<td>As given by Kamper et al, 2014&lt;sup&gt;49&lt;/sup&gt;</td>
<td>13 (NR)</td>
<td>NR (NR)</td>
<td>Pain/ (1)</td>
<td>No</td>
<td>To present a clinically relevant measure to illustrate the benefit of the intervention to patients</td>
<td>No</td>
<td>assessment of risk of bias</td>
<td>No</td>
</tr>
<tr>
<td>Norlund, 2009&lt;sup&gt;49&lt;/sup&gt; the efficacy of IPT</td>
<td>Chronic low back pain lasting more than 4 wks/NR</td>
<td>The definition of multidisciplinary interventions included studies involving two or more healthcare disciplines</td>
<td>7 (1450)</td>
<td>NR (51)</td>
<td>Return to work/ (1)</td>
<td>No</td>
<td>No</td>
<td>Four studies were of moderate quality and three studies were of low quality</td>
<td>No</td>
<td>Critically low</td>
</tr>
</tbody>
</table>


Abbreviations: ACR, American College of Rheumatology; GRADE, Grades of Recommendation, Assessment, Development, and Evaluation; IPT, Interdisciplinary Pain Treatment; mo, months; NR, Not reported; OMERACT, Outcome Measures in Rheumatology Clinical Trials; wks, weeks.
<table>
<thead>
<tr>
<th>Author, Year (Ref)/Objectives</th>
<th>Pain Condition Treated (Definition)/Setting</th>
<th>IPT (Definition)</th>
<th>No. of Included Studies (Total Sample Size)</th>
<th>Mean Age or Age Range, Years (%Female)</th>
<th>Outcomes, (No. of Reported Outcomes)</th>
<th>Distinction Between Primary and Secondary</th>
<th>Specific Strategy Authors Used to Select Outcomes</th>
<th>Risk of Bias Assessment in Included Studies</th>
<th>Evaluation of the Evidence</th>
<th>AMSTAR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martinez-Calderon, 2020&lt;sup&gt;1&lt;/sup&gt; the effectiveness of conservative and surgical interventions</td>
<td>Chronic low back pain (pain persisting for 3 mo or more according to the multidimensional diagnostic criteria for chronic pain by Dworkin et al, 2016) /NR</td>
<td>A minimum of two or more healthcare disciplines (physical, psychological or social)</td>
<td>17 (2017)</td>
<td>35 to 72 (59)</td>
<td>Kinesiophobia, fear-avoidance beliefs, fear of falling/ (3)</td>
<td>No</td>
<td>Fear becomes maladaptive</td>
<td>Half had low risk of bias</td>
<td>GRADE</td>
<td>Moderate</td>
</tr>
<tr>
<td>Martinez-Calderon, 2021&lt;sup&gt;1&lt;/sup&gt; the effectiveness of different interventions</td>
<td>Fibromyalgia syndrome (according to the IASP classification of chronic pain for the ICD-11th edition) /NR</td>
<td>A set of different therapeutic strategies to manage pain</td>
<td>2 (600)</td>
<td>39 to 56 (88)</td>
<td>Fear-avoidance beliefs, fear of movement, fear of pain, kinesiophobia, and pain-related anxiety/ (5)</td>
<td>No</td>
<td>Fear is a key factor in management of chronic musculoskeletal pain</td>
<td>Serious high risk of bias in the majority of included studies</td>
<td>GRADE</td>
<td>Moderate</td>
</tr>
<tr>
<td>Scascighini, 2008&lt;sup&gt;1&lt;/sup&gt; the effectiveness of IPT</td>
<td>Chronic non-specific musculoskeletal pain (eg, chronic low back or back pain, fibromyalgia, no definition is given) / Outpatients and inpatients</td>
<td>At least three of the following categories: psychotherapy, physiotherapy, relaxation techniques, medical treatment or patient education, vocational therapy</td>
<td>27 (2407)</td>
<td>NR (NR)</td>
<td>Pain, emotional strain, quality of life, disability, coping, physical capacity, return to work, sick leave, use of medications, use of the healthcare system, pain behaviour or subjective overall success/ (12)</td>
<td>Yes</td>
<td></td>
<td>Serious high risk of bias in the majority of included studies</td>
<td>The higher effectiveness had to be demonstrated in at least two of the five primary outcomes or in at least one of the primary and two of the secondary outcomes and GRADE</td>
<td>Best evidence synthesis approach by van Tulder criteria (1997)&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td>Guzmán, 2002&lt;sup&gt;1&lt;/sup&gt; the effectiveness of IPT</td>
<td>Chronic low back pain (pain in the lumbar and/or gluteal region with or without radiation to the lower extremities of more than 3 mo) / Pain clinics or hospitals</td>
<td>A minimum of one physical and one of the other dimensions (psychological or social/occupational)</td>
<td>12 (1964)</td>
<td>NR (NR)</td>
<td>Pain severity, global improvement, functional status, quality of life, and employment status/ (5)</td>
<td>No</td>
<td>No</td>
<td>Only three studies had low risk of bias</td>
<td>GRADE</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Author, Year (Ref)/ Objectives</th>
<th>Pain Condition Treated (Definition)/ Setting</th>
<th>IPT (Definition)</th>
<th>No. of Included Studies (Total Sample Size)</th>
<th>Mean Age or Age Range, Years (%Female)</th>
<th>Outcomes, (No. of Reported Outcomes)</th>
<th>Distinction Between Primary and Secondary</th>
<th>Specific Strategy Authors Used to Select Outcomes</th>
<th>Risk of Bias Assessment in Included Studies</th>
<th>Evaluation of the Evidence</th>
<th>AMSTAR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBU-report, 2010(^{19})/ the effectiveness of rehabilitation approaches</td>
<td>Chronic non-cancer pain (pain persisting for 3 mo or more)</td>
<td>A minimum of two or more healthcare disciplines (physical, psychological, or social)</td>
<td>46 (6213)</td>
<td>NR (NR)</td>
<td>Pain, disability or functional status, psychological and cognitive function (depression, anxiety, fear avoidance, coping strategies), generic health or quality of life, work status (return to work, sick leave), healthcare service utilisation, global improvement, sleep quality, medication, pain behaviour, and adverse events/ (14)</td>
<td>Yes</td>
<td>IMMPACT</td>
<td>NR</td>
<td>GRADE</td>
<td>Low</td>
</tr>
<tr>
<td>van Geen, 2007(^{25})/ the effectiveness of IPT</td>
<td>Chronic low back pain (pain persisting for 12 wks or more)/ Outpatients and inpatients</td>
<td>The involvement of several disciplines, such as psychologists, physiotherapists, occupational therapists, and/or medical specialists</td>
<td>10 (1958)</td>
<td>NR (NR)</td>
<td>Work participation, experienced pain, functional status, and quality of life/ (4)</td>
<td>No</td>
<td>No</td>
<td>Half had low risk of bias</td>
<td>More than 50% of the high-quality studies showed effectiveness on the outcome studied</td>
<td>Low</td>
</tr>
<tr>
<td>SBU-report, 2006a, b(^{17})/ the effectiveness of rehabilitation approaches</td>
<td>Chronic non-cancer pain (pain persisting for 3 mo or more)/ NR</td>
<td>A minimum of two or more healthcare disciplines (physical, psychological, or social)</td>
<td>46 (6213)</td>
<td>NR (67)</td>
<td>Pain, disability or functional status, work status (return to work, sick leave), generic health or quality of life, healthcare service utilisation, global improvement, psychological and cognitive function (depression, anxiety, fear avoidance, coping strategies)/ (10)</td>
<td>No</td>
<td>No</td>
<td>NR</td>
<td>Oxford CEMB levels of evidence</td>
<td>Low</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Description</td>
<td>Methodology</td>
<td>Sample Size</td>
<td>Follow-Up</td>
<td>Outcomes</td>
<td>Risk of Bias</td>
<td>Notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garschagen, 2015</td>
<td>2015</td>
<td>Effectiveness of IPT and the need for spiritual care</td>
<td>Chronic non-cancer pain (defined as pain persisting for 3 mo or more)</td>
<td>14 (2672)</td>
<td>NR (NR)</td>
<td>Coping skills, pain, emotional distress, well-being and/or health, employment, and functioning</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waterschoot, 2014</td>
<td>2014</td>
<td>Effectiveness of IPT</td>
<td>Chronic low back pain (disabling non-specific CLBP for at least 3 mo)</td>
<td>18 (3430)</td>
<td>38 to 52 (37 to 100)</td>
<td>Disability, work participation, and quality of life</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nielson, 2001</td>
<td>2001</td>
<td>Effectiveness of IPT and unimodal approaches</td>
<td>Chronic pain (no definition given)</td>
<td>21 (NR)</td>
<td>NR (NR)</td>
<td>Pain, function, and work outcomes</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**


**Abbreviations:**

CEMB, The Centre for Evidence-Based Medicine; CLBP, chronic low back pain; GRADE, Grades of Recommendation, Assessment Development, and Evaluation; IASP, International Association for the Study of Pain; ICD-11, International Classification of Diseases, 11th edition; IMMPACT, Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials; IPT, Interdisciplinary Pain Treatment; mo, months; NR, Not reported; wks, weeks.
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Pain Intensity</th>
<th>Pain Frequency</th>
<th>Physical Activity</th>
<th>Emotional Well-Being</th>
<th>Satisfaction with Social Roles and Activities</th>
<th>Productivity (Paid and Unpaid, at Home and at Work, Inclusive Presentism, and Absenteeism)</th>
<th>Health-Related Quality of Life</th>
<th>Patient’s Perception of Treatment Goal Achievement</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systematic reviews with meta-analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casey, 2020</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>2/8</td>
</tr>
<tr>
<td>Martinez-Calderon, 2020</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>1/8</td>
</tr>
<tr>
<td>Papadopoulou, 2016</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>1/8</td>
</tr>
<tr>
<td>Kamper, 2014</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>2/8</td>
</tr>
<tr>
<td>van Middelkoop, 2011</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>7/8</td>
</tr>
<tr>
<td>Häuser, 2009</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>4/8</td>
</tr>
<tr>
<td>Gianola, 2018</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>4/8</td>
</tr>
<tr>
<td>Norlund, 2009</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>1/8</td>
</tr>
<tr>
<td><strong>Systematic reviews with narrative synthesis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martinez-Calderon, 2021</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>1/8</td>
</tr>
<tr>
<td>Scascighini, 2008</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>7/8</td>
</tr>
<tr>
<td>Guzmán, 2002</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>5/8</td>
</tr>
<tr>
<td>SBU-report, 2010</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>7/8</td>
</tr>
<tr>
<td>van Geen, 2007</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>4/8</td>
</tr>
<tr>
<td>SBU-report, 2006a,b</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>5/8</td>
</tr>
<tr>
<td>Garschagen, 2015</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>3/8</td>
</tr>
<tr>
<td>Waterschoot, 2014</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>3/8</td>
</tr>
<tr>
<td>Nielsen, 2001</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>5/8</td>
</tr>
</tbody>
</table>

**Abbreviations:** Y, Yes; N, No.
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>IMMPACT Outcome Domains</th>
<th>PROMIS Outcome Domains</th>
<th>Score</th>
<th>Physical Health (Including the Core Health Outcome Domains of Symptoms and Function)</th>
<th>Mental Health (Including the Core Health Outcome Domains of Affect, Behaviour, and Cognition)</th>
<th>Social Health (Including the Core Health Outcome Domains of Relationships and Function)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic reviews with meta-analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casey, 2020</td>
<td>Y Y N</td>
<td>N N N</td>
<td>2/6</td>
<td>Y N N</td>
<td>N</td>
<td></td>
<td>1/3</td>
</tr>
<tr>
<td>Martinez-Calderon, 2020</td>
<td>N Y N</td>
<td>N N N</td>
<td>1/6</td>
<td>Y N N</td>
<td>N</td>
<td></td>
<td>1/3</td>
</tr>
<tr>
<td>Papadopoulos, 2016</td>
<td>Y N N</td>
<td>N N N</td>
<td>1/6</td>
<td>Y N N</td>
<td>N</td>
<td></td>
<td>1/3</td>
</tr>
<tr>
<td>Kamper, 2014</td>
<td>Y Y N</td>
<td>N N N</td>
<td>2/6</td>
<td>Y N Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>van Middelkoop, 2011</td>
<td>Y Y Y</td>
<td>Y Y Y</td>
<td>6/6</td>
<td>Y Y Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>Häuser, 2009</td>
<td>Y Y N</td>
<td>N N Y</td>
<td>3/6</td>
<td>Y N Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>Gianola, 2018</td>
<td>Y Y Y</td>
<td>Y N N</td>
<td>3/6</td>
<td>Y Y Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>Norlund, 2009</td>
<td>N Y N</td>
<td>N N N</td>
<td>1/6</td>
<td>N N Y</td>
<td>Y</td>
<td>Y</td>
<td>1/3</td>
</tr>
<tr>
<td>Systematic reviews with narrative synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martinez-Calderon, 2021</td>
<td>N N Y</td>
<td>Y N N</td>
<td>1/6</td>
<td>N Y N</td>
<td>N</td>
<td></td>
<td>1/3</td>
</tr>
<tr>
<td>Scascighini, 2008</td>
<td>Y Y Y</td>
<td>Y N N</td>
<td>5/6</td>
<td>Y Y Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>Guzmán, 2002</td>
<td>Y Y N</td>
<td>N Y N</td>
<td>3/6</td>
<td>Y N Y</td>
<td>Y</td>
<td>N</td>
<td>2/3</td>
</tr>
<tr>
<td>SBU-report, 2010</td>
<td>Y Y Y</td>
<td>Y Y Y</td>
<td>6/6</td>
<td>Y Y Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>van Geen, 2007</td>
<td>Y Y N</td>
<td>N N N</td>
<td>3/6</td>
<td>Y N Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>SBU-report, 2006a,b</td>
<td>Y Y Y</td>
<td>Y N N</td>
<td>4/6</td>
<td>Y Y Y</td>
<td>Y</td>
<td>Y</td>
<td>3/3</td>
</tr>
<tr>
<td>Garschagen, 2015</td>
<td>N Y Y</td>
<td>Y N N</td>
<td>2/6</td>
<td>Y Y N</td>
<td>N</td>
<td>2/3</td>
<td></td>
</tr>
<tr>
<td>Waterschoot, 2014</td>
<td>N Y Y</td>
<td>Y N N</td>
<td>2/6</td>
<td>Y N Y</td>
<td>Y</td>
<td>2/3</td>
<td></td>
</tr>
<tr>
<td>Nielson, 2001</td>
<td>Y Y Y</td>
<td>Y N N</td>
<td>3/6</td>
<td>Y Y Y</td>
<td>Y</td>
<td>3/3</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** Y, Yes; N, No.
in one of the primary and two of the secondary outcomes” and “effectiveness was judged to be robust if more than 50% of the high-quality studies showed effectiveness on the outcome studied” (Tables 1 and 2).

Authors in three SRs, did not report any potential sources of conflict of interest, including any funding they received for conducting their review. Fifteen SRs, assessed the methodological quality of included studies, but less than half clearly stated the tool used and none used the risk of bias of included RCTs as an inclusion criterion. The most common checklists were Cochrane Risk of Bias Tool and van Tulder-11 criteria. The random effect models with Der Simonian and Laird variance estimator were mainly used as statistical synthesis of data in the eight SRs with meta-analysis. All the models used the $I^2$ Index as an indicator of heterogeneity. Subgroup or sensitivity analysis was done by six meta-analytic SRs, but meta-regression analysis was done using only one meta-analytic SR.

**Discussion**

**Summary of Main Results**

This systematic overview showed wide-ranging disparity in reported outcomes and applied outcome domains in SRs evaluating IPT interventions for chronic non-cancer pain conditions. This finding was based on the 18 published SRs with data on more than 49000 people in 356 primary RCTs. Most RCTs reported nociplastic pain (eg, fibromyalgia, low back pain), and no SR focused specifically on neuropathic pain. The duration of IPT varied across the SRs. Primary or secondary outcomes and the rationale of selecting outcome domains were not always stated. Compared to IMMPACT or VAPAIN guidelines, more SRs followed PROMIS to assess outcome domains. However, by counting the number of reported outcome domains according to the three guidelines, we found a positive correlation of VAPAIN with IMMPACT.
and PROMIS, indicating an intercorrelation between the outcome domains included per standardised statement/guideline. A variety of methodological characteristics of SRs are also notable, which affects the quality of the reporting evidence.

**Common Reported Outcomes in SRs Included IPT**

Apart from pain as the most common reported outcome (72%) in all included SRs, disability or functional status (61%) and working status (61%) were also frequently evaluated as outcomes after IPT. Only half of the SRs reported psychological well-being and quality of life. The results reflect the complexity of chronic pain in a biopsychosocial context, and the outcomes after IPT are supposed to include a variety of physical, mental, and social aspects. However, it is uncertain which outcomes are prioritised. For example, whether pain intensity should be included as a major outcome of IPT is often debatable. In addition, SRs did not consider outcomes of adverse events or harms. Lack of specificity on such outcome domains may also bias SR conclusions.

Reporting many outcomes in SRs did not increase the quality of the SRs as a negative correlation was found between AMSTAR2 rating and total number of outcomes. Based on these results, we believe the guidelines for complex intervention for conducting a SR on pain should be re-evaluated. Recently, a new PRISMA extension, known as PRISMA-CI statement and checklist, has been developed to address the important reporting gaps relative to complex interventions in healthcare.

**Why is It Difficult to Report Standardised Outcomes of IPT?**

We need to be aware that the evaluation of complex interventions such as IPT is not clear cut and definitions of a positive outcome of an IPT trial vary in different publications. Although VAPAIN, IMMPACT, and PROMIS are established recommendations for assessing IPT outcomes, few SRs applied them to categorise their outcome domains. A main consequence might be a decreased quality of SRs since a negative correlation between AMSTAR2 and PROMIS was noted. A possible reason of not using these guidelines can be overlooking of the complexity of IPT. In this systematic overview, we found no SR clearly presented the complexity of IPT and considered the intercorrelations between the selected outcomes. For example, outcome measures evaluated separately in one SR may be problematic as the outcomes are most likely to be intercorrelated.

**Methodological Characteristics Affect the Quality of the Evidence**

The approach GRADE for evidence ratings was used in many but not all included SRs. However, since IPT is a complex intervention with high levels of heterogeneity and indirectness assessment, using only GRADE may not adequately describe the evidence base. The latter is supported by Movsisyan and colleagues; they found that the outcomes of complex interventions were more likely to be rated as “very low” quality of evidence compared with those of simple interventions (37.5% vs 9.1% for the primary benefit outcomes). New methods for grading the evidence such as threshold analysis in guideline development have been recently used. We also found diversities in methodological quality assessments of included RCTs and even more often did not state the tool they used in assessment as well as did not use risk of bias of RCTs as an inclusion criterion. With respect to included meta-analytic SRs, we found that subgroup or sensitivity analysis and meta-regression analysis were seldom conducted. Such methodological shortcomings may inflate the quality of the reported evidence.

**Our Suggestions on Handling the Multi-Correlated Outcomes of IPT in SRs**

First, clear definitions of positive outcomes from IPT need to be presented in each SR. RCTs’ different definitions of positive outcomes from IPT should be addressed before an SR is conducted. For example, some SRs treated an outcome of a RCT as a positive outcome when a majority of outcomes were significantly better than for the control intervention. In another SR, the authors predetermined primary and secondary outcomes and what was necessary to classify an intervention as positive before reviewing the RCTs. The recommended guidelines may help categorise the variables into separate domains and minimise the risk of multivariate correlations. Furthermore, advanced graphical approaches such as harvest plots or bubble plots can be useful tools for illustrating and synthesising the matrix of the negative effects, no effects, and positive effects across RCTs in a SR with complex interventions.
Second, intercorrelations between reported outcomes for different RCTs should be considered. One approach is to use the suggested outcome guidelines along with preliminary responder criteria based on several variables per outcome. For example, the evaluation of IPTs for patients with pain in fibromyalgia may be incorporated using both OMERACT core domains and preliminary responder criteria that best favour IPT over control.\textsuperscript{47,73,74} Another approach is to use appropriate statistical methods to handle the intercorrelations. Multivariate methods that can handle intercorrelated outcomes in one analysis have been presented.\textsuperscript{75,76} We have recently suggested how simultaneous goals can be handled using scores from principal component analysis (PCA) in RCTs and observational studies.\textsuperscript{15} Other methods can obtain non-dependent effect sizes and include rationales for selecting effect sizes such as a random selection of effect sizes or averaging them to minimise their interrelations.\textsuperscript{76} For example, when the measurement of pain is evaluated with two or more instruments (eg, VAS and NRS), then it may be more feasible to average the effect size of both instruments into one effect size before meta-analysis of the data.

Third, rather than conducting separate univariate meta-analysis, researchers may consider applying multivariate meta-analysis (MVMA) to provide a framework to address multi-correlated outcomes from IPT.\textsuperscript{77,78} MVMA can simultaneously analyse multiple outcomes of interest and the summary result for each outcome depends on correlated results from related outcomes. However, implementing such an approach requires information about the correlations between the effect sizes, which is rarely reported in original RCTs, as well as knowledge of advance statistical programs like R software.\textsuperscript{76,79} Other meta-analytic strategies include multilevel hierarchical modelling which can show the variation in the true effect sizes across studies or multivariate meta regression which can allow intervention characteristics and mediating effects of intermediate outcomes to be examined together.\textsuperscript{70,76} Finally, mixed treatment comparison meta-analysis methods can also be useful for exploring the different components and combinations of components of IPT as such an approach can examine the effectiveness of a particular component (or combination of components).\textsuperscript{80}

**Strengths and Limitations**

Most importantly, to our knowledge, this is the first overview that addresses disparity in reported outcomes and applies outcome domains in published SRs when evaluating IPT for chronic pain conditions. Using the latest AMSTAR2 checklist, we were able to evaluate the evidence quality of all the included SRs. Despite the complexity of IPT, we mapped, categorised, and compared the reported outcomes according to the VAPAIN statement, IMMPACT, and PROMIS recommendations to evaluate the current evidence of pain rehabilitation processes. This review was designed by experienced reviewers and researchers in this field, guided by a protocol, and carefully stated the difference between the protocol and this review.

This systematic overview has several limitations. First, selection bias cannot be avoided. SRs of RCTs may be associated with risk for bias resulting from an unrepresentative selection of patients and researcher allegiance.\textsuperscript{81} Other study designs, for example, SRs with registry cohorts and observational studies, were not included. Thus, real-world data are lacking. Second, in addition to the varied components of IPT, IPT had great variability in terms of duration of treatment and time of outcome assessment. Whether the reported outcomes from RCTs were related to these factors is not reflected in the present work. Third, the search strategy was limited to three electronic databases and SRs published in English or Swedish. There were no searches of the grey literature, so we might have missed some eligible reviews. However, we searched PROSPERO to identify any ongoing SRs on the topic. One may also argue that our investigation is limited to a small number of included SRs and therefore we may have not captured the entire field. However, we adopted very strict inclusion criteria related to the topic to assess the cutting edge evidence. Finally, the study population considered, ie, different pain conditions and the definition adopted in the SRs for IPT might have influenced the outcomes reported. Given these limitations, future research is needed to improve the framing the selection of research outcomes of IPT.

**Implications for Future Research**

To facilitate evidence synthesis and assessment in complex treatments for chronic pain in every day clinical practice, we need to find a proper way to frame the selection of the research on IPT outcomes. The currently established guidelines for assessing IPT outcomes (VAPAIN, IMMPACT, and PROMIS) are important for categorising outcome domains.
Moreover, researchers who conduct SRs and meta-analyses are supposed to appropriately handle the intercorrelations between outcomes and the interactions between components of complex interventions.71,72 There is also a need to develop clinically accepted ways to apply the established guidelines for reporting outcome domains of IPT. A lack of consensus of evaluating reported outcomes of IPT in SRs makes it difficult to decide whether the evidence from SRs should be applied in clinical practice. A lack of consensus may also confuse health-care providers and policy makers when evaluating the need of tailoring current IPT. Clinical implications will suggest how to report outcomes that capture the information from real-world data to assess the effects of IPT.

**Conclusion**

Currently, there is a wide-ranging disparity in reported outcomes and applied outcome domains in SRs evaluating IPT for chronic pain conditions. According to the definition of complex interventions, IPT has two common characteristics: it has intercorrelated outcomes or mediators and moderators of effect (outcome complexity) and it has multiple components (intervention complexity).82 Given this, we present a menu of SR procedures for addressing sources of complexity when answering questions about the effects of IPT in research synthesis:

- The SRs of IPT should follow the PRISMA-CI extension.
- The inclusion criteria should clearly identify and describe the outcome domains and provide guidance for handling different outcome measurements on the same domain.
- The selection of outcomes should follow the established guidelines such as VAPAIN.
- The scope of included outcomes should address both effectiveness and harms on which strength of the evidence will be graded.
- Quantitative graphical synthesis approaches such as harvest or bubble plots should be adopted to illustrate patterns in results of multiple outcomes. They can also be used as an alternative method of vote counting in SRs when meta-analysis is not feasible.
- Subgroup analysis and meta-regression should be performed to examine how features of the interventions impact effect size. Mixed treatment comparison meta-analysis may also be useful when examining the intervention complexity and the effects of the different intervention components.
- More advanced meta-analytical methods are required that will encourage the exploration and handling of the intercorrelation of the outcomes.76
- There is a clear need for the development and implementation of new methods of grading evidence of IPT and, considering proper framing of the questions, judgements about directness and consistency of evidence and the need for additional contextual and qualitative evidence are needed that provide information about the circumstances when the intervention works best.66
- We believe that our work is the first step to further test our suggestions in future evidence synthesis.

**Abbreviations**

IPT, interdisciplinary pain treatment; SRs, systematic reviews; AMSTAR2, Assessment of systematic Reviews version 2; VAPAIN, Validation and application of a core set of patient-relevant outcome domains; IMMPACT, Initiative on methods, measurements, and pain assessments in clinical trials; PROMIS, Patient-reported outcomes measurement information system.

**Acknowledgments**

This study was supported by grants from County Council of Östergötland (SC-2021, to Huan-Ji Dong), Sweden. The content is solely the responsibility of the authors and does not necessarily represent the official views of County Council of Östergötland.
Disclosure
Björn Gerdle reports grants from Vetenskapsrådet (Swedish research council), outside the submitted work. The authors report no other potential conflicts of interest in relation to this work.

References
17. Swedish Council on Health Technology Assessment (SBU); 2006.


41. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008. doi:10.1136/bmj.j4008


Journal of Pain Research

Publish your work in this journal

The Journal of Pain Research is an international, peer reviewed, open access online journal that welcomes laboratory and clinical findings in the fields of pain research and the prevention and management of pain. Original research, reviews, symposium reports, hypothesis formation and commentaries are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/journal-of-pain-research-journal

Dovepress

Journal of Pain Research 2022:15

Dovepress