ISSN: 2641-1911
DOI: 10.33552/ANN.2026.18.000941
Archives in Neurology & Neuroscience
Iris Publishers

Review Article

RESPONSE - Interdisciplinary Outpatient Multimodal Pain Treatment After Non-Confirmation of Surgical Indication for Persistent Spinal Pain: Real-World Outcomes in 9,217 Patients

Michael A Überall[1*], Thomas Nolte[2], Gerhard HH Müller-Schwefe[3] and Harry Kletzko[4]

[1]*Private Institute of Neurological Sciences – IFNAP, Center of Excellence in Health Care Research, Nürnberg, Germany*

[2]*Center for Outpatient Palliative Care, Wiesbaden, Germany*

[3]*Center for Interdisciplinary Pain and Palliative Care, Göppingen, Germany*

[4]*Integrative Managed Care GmbH, Limburg, Germany*

***Corresponding author:** Michael A. Überall, Private Institute of Neurological Sciences – IFNAP, Center of Excellence in Health Care Research, Nordostpark 51, 90411 Nürnberg, Germany.

Received Date: April 25,2026

Published Date: May 19, 2026

Abstract

Background: Interdisciplinary multimodal pain therapy is recommended for patients with persistent spinal pain who do not respond adequately to unimodal treatments. In many health care systems, however, intensive multimodal programs are predominantly delivered in inpatient settings, and evidence regarding structured intensive outpatient multimodal treatment programs remains comparatively limited. The present study evaluated outcomes of an interdisciplinary outpatient multimodal pain treatment program in patients referred for a second opinion prior to planned spinal surgery.

Methods: This real-world observational study analyzed data from patients treated within a nationwide interdisciplinary outpatient pain treatment network in Germany between January 2014 and December 2025. Patients were referred for interdisciplinary second-opinion assessment prior to a planned spinal surgical intervention within a structured integrated care pathway. Individuals in whom surgery was not confirmed and who subsequently entered an intensive outpatient multimodal treatment program were included in the analysis. Outcomes were assessed using validated patient-reported outcome measures reflecting the biopsychosocial model of chronic pain, including pain intensity, pain-related disability (modified Pain Disability Index; mPDI), functional capacity, psychological distress, health-related quality of life, and pain-related self-efficacy. Treatment- associated changes between baseline and week 3 were analyzed using paired statistical tests and severity-stratified analyses according to the Chronic Pain Grading Scale (CPGS), complemented by multivariate covariance models to evaluate change patterns across outcome domains. The primary outcome was change in pain-related disability after three weeks of treatment.

Results: A total of 9,217 patients entered the outpatient multimodal treatment program and were included in the analysis. Significant improvements were observed across all outcome domains after three weeks of treatment. Pain-related disability, the primary outcome, decreased markedly across all baseline severity strata and was associated with large within-group effect sizes. Clinically meaningful improvements were observed not only in patients with moderate symptom burden but also in those classified as CPGS IV, representing the subgroup with the highest baseline disability. Responder analyses based on predefined absolute and relative improvement thresholds for pain-related disability further supported these findings. Significant improvements were also observed for pain intensity, functional capacity, psychological distress, health-related quality of life, and pain-related self-efficacy. Multivariate analyses demonstrate consistent treatment-associated change patterns across outcome domains and indicate no relevant center-or demography-related effects across participating treatment centers.

Conclusions: In this large real-world cohort of patients with persistent spinal pain referred for interdisciplinary second-opinion assessment prior to planned spinal surgery, intensive interdisciplinary outpatient multimodal pain treatment was associated with substantial short-term improvements across multiple clinically relevant outcome domains. Structured outpatient multimodal programs may therefore represent an effective treatment option even for patients with high baseline disability and may serve as an important component within stepped-care models of pain management.

Keywords: Persistent spinal pain; Multimodal pain therapy; Interdisciplinary pain treatment; Outpatient pain management; Second opinion; Real-world evidence; German Pain e-Registry

Introduction

Persistent spinal pain represents one of the most common causes of long-term disability worldwide and is associated with substantial individual and societal burden [1-3]. Despite the high prevalence of these conditions, the relationship between structural findings in spinal imaging and the severity of pain or functional impairment is often weak [4-6]. As a result, clinical decision-making in patients with persistent spinal pain remains challenging, particularly when surgical interventions are considered.

Current clinical guidelines emphasize the importance of conservative management strategies and highlight the role of interdisciplinary multimodal pain therapy for patients who do not respond adequately to unimodal treatments [7-9]. Multimodal approaches integrate medical, physical, and psychological interventions [10-13]. and aim to address the complex biopsychosocial mechanisms underlying chronic pain [14,15]. While the effectiveness of multimodal pain therapy has been demonstrated in numerous studies, much of the available evidence originates from inpatient treatment programs, and data on intensive outpatient multimodal approaches remain comparatively limited.

In parallel, several healthcare systems have introduced second-opinion procedures prior to planned spine surgery as a strategy to improve the quality of surgical indications and to strengthen guideline- concordant conservative management [16,17]. In routine care, these second-opinion assessments are typically performed by a single specialist who reviews the surgical indication primarily from a biomedical perspective. However, the predominantly mono-disciplinary structure of conventional second-opinion procedures has been critically discussed by pain specialists, as such assessments may remain focused mainly on structural findings and the technical feasibility of surgical interventions rather than on the overall therapeutic benefit for the patient. Against the background of increasing recognition of the biopsychosocial mechanisms underlying persistent spinal pain, pain medicine experts in Germany therefore sought to develop a more comprehensive approach to pre-surgical evaluation. Building on contemporary concepts of multidisciplinary pain management, an interdisciplinary second-opinion model was developed that integrates medical, physiotherapeutic, and psychological expertise in the assessment of patients with persistent spinal pain. This model was implemented within a structured specialized healthcare program in cooperation with statutory health insurance providers and is developed, maintained, and coordinated by the Integrative Managed Care (IMC) GmbH. Since its introduction in 2014, this interdisciplinary assessment model has been offered to insured patients of participating health insurance providers as an additional service complementing routine care. Currently, approximately 30% of the statutory health insurance population in Germany has access to this specialized interdisciplinary second-opinion IMC-program.

The interdisciplinary assessment approach was designed to address the multidimensional nature of persistent spinal pain more comprehensively. Instead of relying on the opinion of a single surgical specialist, the evaluation is conducted by an interdisciplinary team consisting of experienced pain physicians, physiotherapists, and psychologists with expertise in conservative pain management. The rationale for this approach is based on the recognition that surgical decision-making in patients with persistent spinal pain cannot be adequately guided by structural findings alone but requires a comprehensive biopsychosocial evaluation of pain mechanisms, functional impairment, psychological factors, and prior treatment trajectories.

Within this interdisciplinary framework, the team jointly evaluates the overall clinical situation and develops treatment recommendations based on the integrated perspectives of the participating disciplines. These recommendations are subsequently discussed with the patient, and the final therapeutic strategy is determined within a structured shared decision-making process that integrates the interdisciplinary clinical assessment with the patient's individual preferences, expectations, and treatment goals [18].

Patients referred for second-opinion assessment prior to planned spinal surgery often represent a highly complex and severely affected subgroup of individuals with persistent spinal pain who have exhausted conventional unimodal treatment options. Many of these patients have experienced long- standing pain, substantial functional impairment, and considerable psychosocial burden. In routine care, they have frequently undergone multiple unimodal treatment approaches—including pharmacological, physical, and interventional therapies—without achieving satisfactory and sustained improvement. For a subset of these patients, the limitations of conventional unimodal treatment strategies may lead to increasing therapeutic escalation, including consideration of surgical interventions, even in clinical situations where the expected benefit of surgery remains uncertain. This situation reflects the well-recognized challenge that persistent spinal pain often arises from complex interactions between biological, psychological, and social factors that cannot be adequatel addressed through isolated treatment modalities. Within the interdisciplinary second-opinion framework described before, patients in whom the surgical indication is not confirmed are therefore offered an alternative treatment pathway. In cases characterized by particularly high levels of pain- related disability, psychosocial complexity, or extensive prior treatment failure, patients may enter a structured interdisciplinary outpatient multimodal pain treatment program. This program provides an intensive, highly individualized combination of medical, physiotherapeutic, and psychological interventions and thus represents an outpatient treatment approach that in many respects corresponds to the intensity and multidimensional structure traditionally associated with inpatient multimodal pain therapy.

Objectives

The primary objective of the present study was to evaluate the real-world effectiveness of a structured interdisciplinary outpatient multimodal pain treatment program in a large consecutive cohort of patients with persistent spinal pain following interdisciplinary second-opinion assessment in which the surgical indication was not confirmed. We specifically aimed to determine whether clinically meaningful reductions in pain-related disability can be achieved

across predefined baseline severity strata based on the Chronic Pain Grading Scale

(CPGS), [19,20]. a widely used instrument for stratifying chronic pain severity in clinical and epidemiological studies, [21,22]. including patients classified as highly disabled (CPGS IV). In addition, we examined whether baseline severity modifies treatment response and explored treatment-associated changes in secondary patient-reported outcomes across multiple biopsychosocial domains.

Methods

Study Design

The present investigation represents a retrospective cohort analysis of prospectively collected real-world routine care data derived from a structured nationwide interdisciplinary pain treatment network in Germany (Integrative Managed Care GmbH, IMC). The study period extended from January 2014 to December 2025. The underlying care model is embedded within an integrated care framework according to §140a SGB V and includes a standardized interdisciplinary second-opinion procedure for patients with persistent spinal pain who were referred prior to planned spine surgery. The second-opinion assessments were conducted by specialized interdisciplinary teams according to predefined structural and procedural standards. Based on the outcome of this assessment, patients were allocated to one of three predefined care pathways: confirmation of surgical indication, referral to optimized standard outpatient care, or enrollment in an intensive interdisciplinary multimodal outpatient treatment program. Allocation was determined exclusively through interdisciplinary clinical evaluation following predefined care algorithms; no randomization procedures were applied. This study did not interfere with clinical decision-making processes at any stage. The present analysis focuses on the cohort of patients in whom the surgical indication was not confirmed and who subsequently initiated the intensive interdisciplinary multimodal outpatient treatment program. Patients with confirmed surgical indication and those referred to optimized standard care were excluded from the analytical dataset. Accordingly, the study represents a real-world effectiveness evaluation of an intensive interdisciplinary multimodal outpatient intervention embedded within a structured clinical decision pathway following interdisciplinary second-opinion assessment. Outcome assessments were performed at two predefined time points: baseline prior to initiation of the outpatient treatment program (T0) and after completion of the three-week intervention (T3). Data were documented within the German Pain e-Registry (GPeR), a standardized electronic documentation platform used in routine pain care.

Research Questions

The primary objective of this study was to evaluate the effectiveness of an intensive interdisciplinary multimodal outpatient pain treatment program in patients with persistent spinal pain who had been referred to a second opinion prior to planned spine surgery and in whom the surgical indication was not confirmed.

The primary outcome was short-term change in pain-related disability following completion of the three-week outpatient program, assessed using the modified Pain Disability Index (mPDI).

Secondary objectives were to examine treatment-related changes across multiple pain-related domains, including pain intensity, functional capacity, psychological distress, fear-avoidance beliefs, neuropathic pain features, health-related quality of life, pain-related quality-of-life impairment, pain self-efficacy, and habitual well-being.

An additional objective was to examine whether patients with higher baseline pain severity and chronicity derive clinically meaningful benefit from an outpatient interdisciplinary multimodal treatment program. To address this question, outcomes were evaluated across predefined severity strata based on the Chronic Pain Grading Scale (CPGS), including patients classified as highly disabled (CPGS IV).

Methods

Setting and Data Source

The study was conducted within a nationwide network of specialized interdisciplinary pain centers operating under the umbrella of Integrative Managed Care GmbH (IMC), Germany. During the observation period, a total of 48 centers participated in the network, although not all centers were active throughout the entire study timeframe.

All participating centers fulfilled predefined structural and organizational criteria for interdisciplinary pain assessment and multimodal treatment delivery. These criteria included the coordinated involvement of physicians specialized in pain medicine, physiotherapists, and psychologists working within an interdisciplinary treatment framework.

Clinical data were derived from routine documentation within the German Pain e-Registry (GPeR), a structured electronic documentation system implemented across the network. The GPeR captures standardized clinical variables and patient-reported outcome measures at predefined assessment time points as part of routine care processes.

Data entry within the GPeR is performed contemporaneously during clinical encounters. The system incorporates automated plausibility checks and completeness controls at the time of entry in order to ensure formal correctness and internal consistency of documentation.

The assessment battery used at baseline (T0) and after completion of the three-week intervention (T3) was predefined and aligned with the standards of the German Pain Questionnaire and its validated supplementary instruments. All instruments employed in the present study were scientifically validated and are recommended for routine outpatient pain assessment by the relevant German pain societies and the German Pain League as responsible national pain patient organization.

All data were stored in pseudonymized form within the registry infrastructure. For the present analysis, fully anonymized datasets were extracted after completion of the observation period. No

additional data collection was performed for research purposes.

Eligibility Criteria and Patient Flow

All patients documented within the registry during the study period had sought an interdisciplinary second opinion prior to planned spine surgery due to persistent spinal pain. Eligibility for the present analysis required completion of the structured interdisciplinary second-opinion assessment and subsequent allocation to the intensive interdisciplinary multimodal outpatient treatment program following non-confirmation of surgical indication.

Patients in whom the surgical indication was confirmed during the second-opinion process were not eligible for inclusion in the present study. Likewise, patients referred to optimized standard outpatient care after non-confirmation of surgical indication were excluded from the analytical cohort.

The analytical dataset comprised all consecutive patients allocated to the intensive interdisciplinary multimodal outpatient program within the predefined observation period.

Baseline (T0) assessment was performed prior to initiation of the outpatient program. Follow-up assessment (T3) was conducted upon completion of the 3-week intervention. As participation in the integrated care program required formal consent to structured documentation procedures, complete baseline and week-3 outcome data were available for all patients included in the analysis. No patients were excluded due to incomplete outcome documentation.

Sample size considerations

Because the present study represents a real-world evaluation of consecutively treated patients within a predefined integrated care pathway, no formal a priori sample size calculation was performed. The cohort size was determined by the total number of eligible patients treated during the predefined observation period.

Given the observational design and the continuous nature of the primary endpoint, precision of effect estimates increases with sample size, resulting in narrow confidence intervals for mean changes and subgroup comparisons in sufficiently large cohorts. The available dataset was therefore considered adequate for the prespecified primary, subgroup, and multivariable analyses.

Intervention

Patients allocated to the intensive interdisciplinary multimodal outpatient program participated in a structured 3-week treatment course comprising 15 working days.

The program followed predefined temporal and organizational frameworks specifying treatment time indows and minimum weekly treatment volumes across medical, physiotherapeutic, and psychological domains. Active patient-directed treatment sessions were scheduled on alternating days (typically Monday, Wednesday, and Friday), with approximately 3–4 hours of direct therapeutic contact per treatment day. On the remaining weekdays (Tuesday and Thursday), structured interdisciplinary team activities were conducted, including review of patient-reported outcomes, case conferences, treatment planning, and dynamic adjustment of therapeutic strategies based on individual patient progress.

Although the program operated within a clearly defined structural framework, therapeutic content was not delivered according to a rigid standardized protocol. Instead, treatment was highly individualized and adapted to the specific clinical presentation, functional limitations, psychosocial context, and therapeutic needs of each patient. Interdisciplinarity and multimodality were core principles guiding treatment implementation.

Therapeutic elements were delivered using a combination of individual and group-based formats. The proportion and sequencing of these formats were determined on a case-by-case basis in order to ensure needs-oriented and patient-centered care while maintaining interdisciplinary coordination.

Core treatment domains included medical pain management and medication optimization, activation-oriented physiotherapy and graded functional restoration, cognitively and behaviorally informed psychological interventions targeting maladaptive coping and fear-avoidance beliefs, and structured patient education aimed at enhancing self-management and pain-related self-efficacy.

All therapeutic components were delivered in accordance with recommendations of the German Federal Joint Committee and current national and international evidence-based guidelines for the management of chronic spinal pain.

In addition to baseline (T0) and end-of-treatment (T3) assessments, standardized interim documentation was performed after week 1 and week 2 to monitor clinical progress and support adaptive treatment planning. Interim assessments were shorter than the full baseline and final assessment battery but followed predefined documentation standards within the registry system.

Within the integrated care framework, reimbursement for the interdisciplinary treatment program followed a bundled payment model rather than a procedure-based fee-for-service structure. This approach allows flexible, patient-centered allocation of therapeutic interventions across disciplines and supports outcome-oriented interdisciplinary care delivery.

Outcome Measures

To comprehensively evaluate treatment-associated changes across the multidimensional nature of persistent spinal pain, a battery of validated patient-reported outcome measures was applied. The selected instruments cover key domains of the biopsychosocial pain model, including pain intensity, pain-related disability, functional capacity, psychological distress, health-related quality of life, fear- avoidance beliefs, self-efficacy, neuropathic pain characteristics, and general well-being. All instruments used in the present study have been scientifically validated and are recommended for comprehensive biopsychosocial pain assessment by the relevant German pain societies and the German Pain League as the national patient organization. Outcome measures were assessed at baseline prior to initiation of the outpatient treatment program (T0) and after completion of the three-week intervention (T3).

The primary outcome of the present study was change in pain-related disability from baseline (T0) to completion of the 3-week outpatient program (T3), assessed using the modified Pain Disability Index (mPDI) [23-25]. The mPDI is a widely used and validated instrument for assessing pain-related disability in patients with chronic pain conditions and assesses pain-related interference across seven domains of daily life (family/home responsibilities, recreation, social activity, occupation, self-care, sleep and life-enjoyment). Each domain is rated by the patient using a 100-mm visual analogue scale (VAS), with higher values indicating greater pain-related interference. For the present analysis, the mPDI score was calculated as the mean of the seven domain-specific VAS ratings and normalized to a 0–100 scale, with higher scores indicating greater pain-related disability and mPDI scores ≥50 indicative for severe pain-related disability. Treatment-related improvement was operationalized as either an absolute reduction of at least 20 mm on the VAS (the minimal clinical important difference; MCID) or a relative reduction of at least 50% compared with baseline values [26,27].

Pain intensity was assessed using 24-hour ratings of least pain intensity (LPI), average pain intensity (API), and highest pain intensity (HPI) [28]. All pain intensity measures were recorded using 100 mm visual analogue scales (VAS), ranging from 0 mm ("no pain") to 100 mm ("worst imaginable pain"), with higher values indicating greater pain intensity [29,30]. For subgroup and sensitivity analyses, clinically relevant thresholds were defined as LPI ≥ 30 mm, API ≥ 50 mm, and HPI ≥ 80 mm. Treatment-related improvement in pain intensity was operationalized as either an absolute reduction of at least 20 mm on the VAS (the minimal clinical important difference; MCID) or a relative reduction of at least 50% compared with baseline values [26,27,31].

Functional impairment due to back pain was assessed using the Hannover Functional Ability Questionnaire (HFAQ), a validated instrument measuring back pain–related limitations in everyday physical activities [32,33]. The questionnaire consists of 12 items addressing common functional tasks of daily living. Responses are transformed into a summary score ranging from 0 to 100, with higher scores indicating better functional ability and lower scores reflecting greater functional limitation. Based on established reference data, a score below 65 was considered indicative of clinically relevant functional impairment [34].

Fear-avoidance beliefs were assessed using the Fear-Avoidance Beliefs Questionnaire (FABQ), a validated instrument measuring pain-related fear and avoidance behavior [35-37]. The questionnaire comprises two subscales addressing fear-avoidance beliefs related to physical activity (FABQ-PA) and work (FABQ-W). The physical activity subscale consists of 4 items (score range 0–24), whereas the work-related subscale comprises 7 items (score range 0–42). Higher scores indicate stronger fear- avoidance beliefs. Based on commonly used clinical thresholds, scores ≥15 on the FABQ-PA and ≥34 on the FABQ-W were considered indicative of elevated risk for persistent disability [38].

Psychological distress was evaluated using the Depression, Anxiety, and Stress Scales – 21 item version (DASS-21), a validated self-report instrument assessing symptoms of depression, anxiety, and stress [39,40]. The questionnaire consists of 21 items rated on a 4-point Likert scale (0–3), with seven items contributing to each of the three subscales: Depression, Anxiety, and Stress. Subscale scores were analyzed separately, with higher scores indicating greater symptom severity. For categorical analyses, cut-offs corresponding to the "strong to extremely severe" symptom range were applied, defined as depression scores ≥11, anxiety scores ≥8, and stress scores ≥13 [41].

Neuropathic pain components were assessed using the 7-item short-version of the painDETECT questionnaire (PDQ-7), a validated screening instrument for identifying neuropathic pain features in patients with chronic pain [42,43]. The PDQ-7 consists of seven sensory symptom items rated on Likert-type scales, yielding a total score ranging from 0 to 35, with higher scores indicating a greater likelihood of neuropathic pain involvement. For categorical stratification, scores ≤12 were interpreted as unlikely neuropathic pain, scores between 13 and 18 as indeterminate, and scores ≥19 as indicative of a likely neuropathic pain component [44].

Health-related quality of life was assessed using the 12-item Veterans RAND Health Survey (VR-12), a validated instrument derived from the SF-12 that captures physical and mental health status [45]. The questionnaire generates two norm-based summary scores: the Physical Component Summary (PCS) and the Mental Component Summary (MCS). Both summary scores are standardized to a population mean of 50 with a standard deviation of 10, with higher values indicating better health-related quality of life. In the present analysis, clinically relevant impairment was defined as PCS scores ≤35 and MCS scores ≤40, reflecting marked limitations in physical and mental health relative to the reference population [46].

General habitual well-being was assessed using the Marburg Questionnaire on Habitual Health Findings (MQHHF), a validated self-report instrument measuring subjective habitual well-being [47]. The questionnaire comprises seven items covering general physical and psychological well-being, resulting in a total score ranging from 0 to 35, with higher scores indicating better perceived well- being. In the present analysis, values ≤10 were interpreted as indicative of pronounced impairment in habitual well-being [48].

Pain-specific self-efficacy was assessed using the Pain Self-Efficacy Questionnaire (PSEQ), a validated 10-item self-report instrument measuring the patient's confidence in performing activities despite pain [49-51]. Items are rated on a 7-point scale, yielding a total score ranging from 0 to 60, with higher scores indicating greater perceived ability to function despite pain. In the present analysis, scores ≤30 were interpreted as indicative of clinically relevant impairment in pain-related self-efficacy.

Pain severity and chronicity were classified using the Chronic Pain Grading Scale (CPGS), a validated instrument designed to categorize chronic pain according to pain intensity and pain-related disability [19,52]. The CPGS combines measures of characteristic pain intensity and pain-related interference with daily activities to assign patients to four hierarchical grades reflecting increasing

levels of chronic pain severity. The instrument has been widely used in epidemiological and clinical pain research because it allows standardized stratification of patients according to clinically meaningful levels of pain-related disability. Characteristic pain intensity is derived from three numerical rating items assessing current pain, average pain, and worst pain during the previous four weeks, while pain- related disability is assessed using three items capturing interference with daily activities, work, and social participation. Based on these dimensions, patients are categorized into four grades: Grade I (low disability–low intensity), Grade II (low disability–high intensity), Grade III (high disability– moderately limiting), and Grade IV (high disability–severely limiting). For the purposes of subgroup analyses in the present study, patients were categorized into three predefined severity strata: CPGS I– II (low disability), CPGS III (moderate disability), and CPGS IV (severe disability) [20]. Grades I and II were combined because both represent pain conditions with low disability, differing primarily in pain intensity rather than functional impairment.

All thresholds and categorical classifications applied in subgroup and sensitivity analyses were derived from published literature, established normative data, or guideline-based clinical recommendations.

Statistical Analysis

All statistical analyses were performed using SPSS (PASW Statistics, Version 18.0; SPSS Inc., Chicago, IL, USA). Continuous variables are reported as mean and standard deviation (SD), whereas categorical variables are presented as absolute and relative frequencies.

Within-group changes from baseline (T0) to week 3 (T3) were analyzed using paired-sample t-tests. Effect sizes for within-group changes were calculated as Cohen's d (=standardized mean difference, SMD) for dependent samples, defined as the mean change divided by the standard deviation of the change scores [53]. Effect sizes were interpreted according to conventional thresholds, with values of approximately 0.2 considered small, 0.5 moderate, and 0.8 or greater large [54].

Baseline differences across Chronic Pain Grading Scale (CPGS) severity strata were examined using one-way analysis of variance (ANOVA) for continuous variables and chi-square tests for categorical variables [55]. Between-group differences in continuous treatment effects were evaluated using analysis of covariance (ANCOVA) models,[56]. with the respective T3 outcome as dependent variable, CPGS group as fixed factor, and the baseline value of the corresponding outcome included as covariate.

To evaluate whether treatment-associated change profiles across multiple outcome domains differed between CPGS severity strata, multivariate analyses of covariance (MANCOVA) were additionally performed [57]. In these models, predefined change scores (T3–T0) of selected outcome domains were entered simultaneously as dependent variables. CPGS severity group was included as the main factor, and treatment center, age and sex were included as a covariates to account for potential clustering across participating centers or demographic effects. Multivariate test statistics were evaluated using Wilks' lambda and Pillai's trace, and corresponding effect sizes were expressed as partial eta-squared (η^2p) [54].

Effect sizes for between-group comparisons were quantified using partial eta-squared (η^2p) for continuous outcomes derived from ANCOVA models. For categorical between-group comparisons, effect sizes were expressed as phi coefficients (for 2×2 tables) or Cramer's V (for larger contingency tables). Effect size magnitudes were interpreted according to established conventions, with η^2p values of approximately 0.01 considered small, 0.06 moderate, and 0.14 large, and phi/Cramer's V values of approximately 0.10 considered small, 0.30 moderate, and 0.50 large [54].

The primary endpoint was defined as change in pain-related disability (mPDI) from T0 to T3. The primary confirmatory hypothesis of the study was that clinically meaningful reductions in pain-related disability can be achieved across all predefined CPGS severity strata, including patients classified as CPGS IV. Responder analyses for the primary endpoint were conducted using predefined absolute and relative improvement criteria for mPDI. Differences in responder proportions between CPGS strata were examined using chi-square tests, with corresponding effect sizes reported as phi or Cramer's as appropriate.

To account for the large number of secondary outcomes, adjustment for multiple testing was performed using the Bonferroni method [58,59]. All statistical tests were two-sided. A p-value <0.05 was considered statistically significant for the primary endpoint, whereas Bonferroni-adjusted thresholds were applied for secondary analyses.

Model assumptions for parametric analyses were examined prior to statistical testing. Normality of residual distributions and homogeneity of variances were assessed using standard diagnostic procedures within SPSS. For multivariate models, homogeneity of covariance matrices was evaluated using Box's M test [60].

Given the large sample size of the study cohort, interpretation of results focused not only on statistical significance but also on the magnitude of observed effect sizes in order to evaluate the clinical relevance of treatment-associated changes.

Confidence intervals were calculated but not routinely reported in the tables in order to maintain readability given the large number of outcomes; however, standardized effect size measures were systematically provided to allow assessment of clinical relevance beyond statistical significance.

Ethical and Regulatory Considerations

The integrated care program evaluated in the present study was conducted within the framework of contractual integrated care agreements according to §140a SGB V (German Social Code, Book V). Participation in the program was voluntary and required written informed consent prior to initiation of the interdisciplinary second-opinion assessment. Patients provided consent both for participation in the structured care program and for the use of pseudonymized clinical data for healthcare research purposes. This consent was documented in written form and subsequently confirmed electronically at first use of the German Pain e-Registry (GPeR).

Clinical data were stored in pseudonymized form within the registry infrastructure. For the present analysis, fully anonymized datasets were extracted. The investigation represents a retrospective analysis of routinely collected healthcare data and did not involve any additional diagnostic or therapeutic interventions beyond those applied as part of the specialized multimodal care program.

Ethical review and approval were waived because the study represents a secondary analysis of anonymized routine care data collected within a structured integrated care program according to §140a SGB V. Nevertheless, the protection of patient rights and data privacy in the context of the present analysis was reviewed and monitored by the Ethics Board of the German Pain League (Deutsche Schmerzliga).

The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki.

Funding and Governance

The present study received no external public or commercial research funding. The analysis was conducted on behalf of Integrative Managed Care GmbH (IMC GmbH) within the framework of its structured healthcare evaluation activities.

Statistical analyses were performed independently by O.Meany-MDPM GmbH under methodological supervision of the Institute for Neurological Sciences (IFNAP) and the Center of Excellence for Health Care Research, Nürnberg, Germany.

The statistical analysis plan was defined prior to data evaluation. All analyses were conducted independently of the network operator. Interpretation of results and manuscript preparation were performed by the authors, all of whom had full access to the anonymized dataset and critically reviewed and approved the final manuscript.

The project was prospectively registered within the European Medicines Agency (EMA) framework as a healthcare research initiative and documented within the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) registry to ensure transparency of study objectives and methodology (EU PAS identifier: EUPAS1000000952).

RESULTS

Study Population and Patient Flow

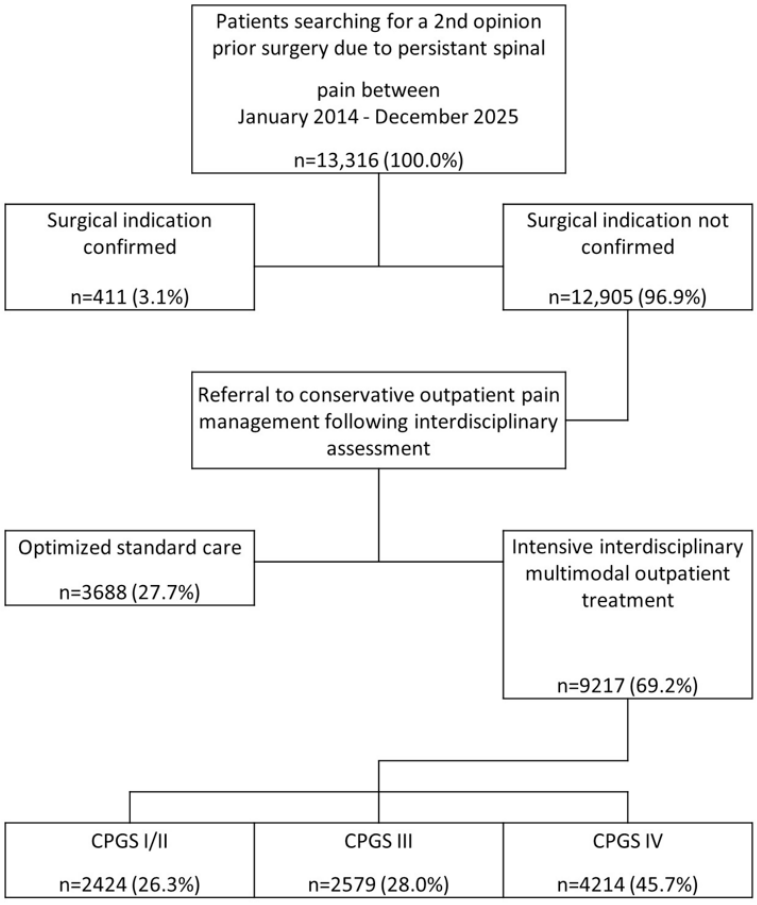

Figure 1: Flow of patients undergoing interdisciplinary second-opinion assessment prior to spine surgery and allocation to treatment pathways. The study population comprises patients referred to the intensive interdisciplinary multimodal outpatient program following non-confirmation of surgical indication.

Between January 2014 and December 2025, a total of 13,316 patients underwent the interdisciplinary second-opinion assessment prior to planned spine surgery within the integrated care network of the IMC (Figure 1). In 411 patients (3.1%), the surgical indication was confirmed by the assessment team. In 3,688 patients (27.7%), surgery was not confirmed and optimized standard outpatient care was recommended. The remaining 9,217 patients (69.2%) – for whom the surgical indication wasn´t confirmed as well – were allocated to the intensive interdisciplinary multimodal outpatient treatment program and constitute the analytical cohort of the present study (Figure 1).

All 9,217 patients completed baseline assessment (T0) prior to initiation of treatment and end-of-program assessment after three weeks (T3). No patients were excluded due to incomplete outcome documentation. According to the Chronic Pain Grading Scale (CPGS), 1,122 patients (12.2%) were classified as grade I, 1,302 (14.1%) as grade II, 2,579 (28.0%) as grade III, and 4,214 (45.7%) as grade IV, indicating that nearly half of the cohort presented with the highest chronic pain severity grade at baseline.

Baseline Characteristics of the Study Cohort

Baseline characteristics of the study cohort across the predefined CPGS severity strata are presented in Tables 1–4. Overall, the cohort represents a highly burdened population of patients with persistent spinal pain who had undergone extensive prior treatment but continued to experience substantial pain, functional impairment, and psychosocial distress at the time of program entry (Tables 1-4).

Table 1: Baseline demographic and clinical characteristics stratified by CPGS severity group.

Cohort	CPGS I/II	CPGS III	CPGS IV
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)
Female gender [n (%)]	1464 (60.4)	1646 (63.8)	2513 (59.6)
Age [years; mean (SD)]	51.0 (14.0)	50.7 (13.5)	50.6 (13.8)
LBP-duration <12 months [n (%)]	1002 (41.3)	1326 (51.4)	2394 (56.8)
1-3 years [n (%)]	528 (21.8)	350 (13.6)	596 (14.1)
>3 years [n (%)]	894 (36.9)	903 (35.0)	1224 (29.0)
Persistent spinal pain syndrome (PSPS): type 1 [n (%)]	1670 (68.9)	1609 (62.4)	2414 (57.3)
type 2 [n (%)]	754 (31.1)	970 (37.6)	1800 (42.7)
LBP-related sickleave in the last 3 months: none [n (%)]	666 (27.5)	126 (4.9)	31 (0.7)
transient [n (%)]	1728 (71.3)	2362 (91.6)	3309 (78.5)
continuously [n (%)]	30 (1.2%)	91 (3.5)	874 (20.7)
Number of LBP-related sickleave days in the last 3 months [n; mean (SD)]	17.2 (24.4)	34.1 (25.5)	62.2 (25.0)
Persistent pain with ...minor fluctuations [n (%)]	851 (35.1)	846 (32.8)	1264 (30.0)
...major fluctuations [n (%)]	967 (39.9)	929 (36.0)	1423 (33.8)
...pain attacks [n (%)]	606 (25.0)	804 (31.2)	1527 (36.2)
Lowest 24-hr. pain intensity [mm VAS; mean (SD)]	17.9 (13.8)	23.1 (16.5)	26.2 (19.6)
LPI≥30 mm VAS [n (%)]	407 (16.8)	720 (27.9)	1471 (34.9)
Average 24-hr. pain intensity [mm VAS; mean (SD)]	42.5 (19.0)	51.8 (18.2)	55.2 (19.3)
API≥50 mm VAS [n (%)]	964 (39.8)	1520 (58.9)	2724 (64.6)
Highest 24-hr. pain intensity [mm VAS; mean (SD)]	66.2 (21.0)	75.0 (18.9)	79.0 (17.1)
HPI≥80 mm VAS [n (%)]	784 (32.3)	1325 (51.4)	2537 (60.2)
PDQ7 pain phenotype score [NRS-35; mean (SD)]	11.9 (7.1)	14.3 (7.7)	14.7 (7.0)
Neuropathic phenotype [n (%)]	477 (19.7)	752 (29.2)	1417 (33.6)
LBP-related disability in daily life [mPDI; mm VAS; mean (SD)]	42.9 (16.9)	56.0 (15.4)	65.1 (14.9)
mPDI≥50 mm VAS [n (%)]	790 (32.6)	1578 (61.2)	3514 (83.4)
LBP-related sleep problems [n (%)]	796 (32.8)	1249 (48.4)	2448 (58.1)
Comorbidities [ICD10 GM 2025; n; mean (SD)]	1.6 (1.7)	1.8 (1.7)	1.8 (1.9)
Two or more comorbidities [n (%)]	1024 (42.2)	1206 (46.8)	1799 (42.7)
Comedication [ATC; mean (SD)]	2.5 (1.8)	2.4 (1.8)	2.5 (1.8)
Two or more comedications [n (%)]	1626 (67.1)	1681 (65.2)	2819 (66.9)

Baseline demographic, clinical, and occupational characteristics of patients undergoing interdisciplinary outpatient multimodal treatment following second-opinion assessment prior to spine surgery. Data are presented as mean (SD) or n (%). SD = standard deviation; CPGS = Chronic Pain Grading Scale; PSPS = Persistent Spinal Pain Syndrome; LBP = low back pain.

Table 2: Baseline health status, psychological burden, and quality-of-life measures across CPGS severity groups.

Cohort	CPGS I/II	CPGS III	CPGS IV
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)
Generall wellbeing (MQHHF; NRS-35; mean (SD)]	15.2 (7.1)	11.6 (7.0)	9.5 (6.4)
MQHHF≤10 [n (%)]	634 (26.2)	1235 (47.9)	2498 (59.3)
Physical quality-of-life [VR12-PCS; mean (SD)]	35.3 (9.5)	30.7 (8.6)	27.4 (8.3)
VR12-PCS≤35 [n (%)]	1051 (43.4)	1785 (69.2)	3475 (82.5)
Mental quality-of-life [VR12-MCS; mean (SD)]	43.2 (11.8)	38.7 (11.4)	38.0 (11.5)
VR12-MCS≤40 [n (%)]	985 (40.6)	1527 (59.2)	2524 (59.9)
Fear avoidance beliefs - work [FABQ-W; NRS-42; mean (SD)]	13.8 (11.6)	18.3 (12.4)	20.4 (11.8)
FABQ-W≥34 [n (%)]	184 (7.6)	368 (14.3)	690 (16.4)
Fear avoidance beliefs - physical activity (FABQ-PA; NRS-24; mean (SD)]	13.1 (6.0)	15.2 (9.3)	15.7 (5.5)
FABQ-PA≥15 [n (%)]	1088 (44.9)	1475 (57.2)	2689 (63.8)
Hannover Functional Questionnaire Backache [HFQB; NRS-100; mean (SD)]	68.5 (23.0)	58.4 (22.7)	50.7 (22.4)
HFQB<65 [n (%)]	815 (33.6)	1512 (58.6)	3066 (72.8)
Pain self efficacy [PSEQ; NRS-60; mean (SD)]	34.7 (15.1)	29.9 (12.7)	26.6 (11.6)
PSEQ≤30 [n (%)]	917 (37.8)	1384 (53.7)	2743 (65.1)
Work-related behaviour and experience patterns: type A [n (%)]	510 (21.0)	534 (20.7)	1049 (24.9)
type B [n (%)]	683 (28.2)	915 (35.5)	1365 (32.4)
type S [n (%)]	1034 (42.7)	1048 (40.6)	1507 (35.8)
type G [n (%)]	197 (8.1)	82 (3.2)	293 (7.0)
Depression [DASS-21 D; NRS-21; mean (SD)]	5.1 (4.9)	6.7 (5.2)	7.4 (5.2)
Strong to extreme depression [DASS-21 D≥11; n (%)]	331 (13.7)	568 (22.0)	1133 (26.9)
Anxiety [DASS-21 A; NRS-21; mean (SD)]	3.7 (4.6)	4.3 (4.9)	4.6 (4.7)
Strong to extreme anxiety [DASS-21 A≥8; n (%)]	370 (15.3)	510 (19.8)	1006 (23.9)
Stress [DASS-21 S; NRS-21; mean (SD)]	7.3 (5.1)	9.1 (5.3)	9.1 (5.2)
Strong to extreme stress [DASS-21 S≥13; n (%)]	444 (18.3)	639 (24.8)	1146 (27.2)

Baseline patient-reported outcome measures (PROMs) reflecting pain phenotype, general well-being, quality of life, fear-avoidance beliefs, functional capacity, psychological distress, and self-efficacy across CPGS severity groups. Continuous variables are presented as mean (SD); categorical thresholds as n (%). SD = standard deviation; VR12 = Veterans RAND 12-Item Health Survey; DASS-21 =Depression Anxiety Stress Scales; PSEQ = Pain Self-Efficacy Questionnaire.

Table 3: Previous interventional and pharmacological treatment exposure stratified by CPGS severity.

Cohort	CPGS I/II	CPGS III	CPGS IV
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)
Previous neural therapy [NT; n (%)]	1578 (65.1)	1911 (74.1)	3455 (82.0)
Previous periradicular Therapy [PRT; n (%)]	775 (32.0)	1048 (40.6)	2103 (49.9)
Previous LBP-related spinal surgery [n (%)]	754 (31.1)	970 (37.6)	1800 (42.7)
Previous LBP-related background drug treatments [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)
Non-steroidal antiinflammatory drugs [NSAIDs; n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)
Nonopioid analgesics [NOA; n (%)]	1885 (77.8)	1945 (75.4)	3261 (77.4)
Mild opioid analgesics [MOA; n (%)]	1889 (77.9)	2017 (81.7)	3298 (78.3)
Strong opioid analgesics [SOA; n (%)]	1106 (45.6)	1095 (42.5)	1921 (45.6)
Cannabinoids [CAN; n (%)]	312 (12.9)	333 (12.9)	613 (14.5)
Muscle relaxants (MREL; n (%)]	1786 (73.7)	1869 (72.5)	3162 (75.0)
Antidepressant drugs (ADD; n (%)]	2085 (86.0)	2178 (84.5)	3586 (85.1)
Antiepileptic drugs [AED; n (%)]	1288 (53.1)	1280 (49.6)	2132 (50.6)
Previous background drug treatments [n; mean (SD)]	8.9 (2.2)	8.8 (2.2)	8.9 (2.3)
Previous rescue drug treatments [n (%)]	1640 (67.7)	1783 (69.1)	2856 (67.8)

History of interventional procedures and background pharmacological treatments prior to initiation of the outpatient multimodal program. Values are presented as n (%) or mean (SD). SD = standard deviation; NSAIDs = non-steroidal anti-inflammatory drugs; MOA = mild opioid analgesics; SOA = strong opioid analgesics; CAN = cannabinoids; AED = antiepileptic drugs; ADD = antidepressant drugs.

Table 4: Current pharmacological treatment at baseline (pre-treatment status).

Cohort	CPGS I/II	CPGS III	CPGS IV
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)
Current background drug treatments [n; mean (SD)]	4.5 (1.5)	4.4 (1.5)	4.5 (1.6)
Non-steroidal antiinflammatory drugs [NSAIDs; n (%)]	1752 (72.3)	1901 (73.7)	3168 (75.2)
Nonopioid analgesics [NOA; n (%)]	1022 (42.2)	1027 (39.8)	1741 (41.3)
Mild opioid analgesics [MOA; n (%)]	1136 (46.9)	1179 (45.7)	1859 (44.1)
Strong opioid analgesics [SOA; n (%)]	628 (25.9)	640 (24.8)	1124 (26.7)
Cannabinoids [CAN; n (%)]	221 (9.1)	241 (9.3)	394 (9.3)
Muscle relaxants (MREL; n (%)]	1786 (73.7)	1867 (72.4)	3162 (75.0)
Antidepressant drugs (ADD; n (%)]	1765 (72.8)	1791 (69.4)	2970 (70.5)
Antiepileptic drugs [AED; n (%)]	607 (25.0)	570 (22.1)	921 (21.9)
Current rescue drug treatments [n (%)]	1170 (48.3)	1288 (49.9)	2020 (47.9)

Ongoing background and rescue pharmacological treatments at baseline before initiation of the 3-week interdisciplinary outpatient program. Values are presented as mean (SD) or n (%). SD = standard deviation; NSAIDs = non-steroidal anti-inflammatory drugs; MOA = mild opioid analgesics; SOA = strong opioid analgesics; CAN = cannabinoids; AED = antiepileptic drugs; ADD = antidepressant drugs.

Across the different domains assessed—including demographic characteristics, pain-related disability, psychological burden, prior treatment exposure, and current pharmacological therapy—consistent gradients toward greater disease burden were observed with increasing CPGS severity.

Demographic and Clinical Characteristics

Baseline demographic and clinical characteristics stratified by Chronic Pain Grading Scale (CPGS) severity strata are summarized in Table 1. Sex distribution and age were comparable across severity strata. Women accounted for approximately 60% of patients in each group, and mean age was similar across groups (approximately 51 years). Pain duration showed moderate variation between severity strata. Pain duration shorter than 12 months was reported more frequently in patients with higher CPGS severity, whereas pain duration exceeding three years was somewhat more common in the CPGS I/II and CPGS III groups. Regarding pain classification, persistent spinal pain syndrome type 1 (PSPS-1) was the predominant phenotype across all severity strata but decreased with increasing severity, whereas the proportion of PSPS-2, reflecting persistent pain following previous spinal surgery, increased correspondingly with higher CPGS severity. Substantial differences between severity groups were observed for work impairment. Absence of sick leave in the preceding three months was common among patients in the CPGS I/II group but rare in patients with CPGS III and nearly absent in CPGS IV. Conversely, continuous sick leave and the number of pain-related sick leave days increased markedly with higher CPGS severity. Pain patterns also differed across severity strata. While persistent pain patterns predominate in patients with lower CPGS grades, recurrent pain attacks were more frequently reported in patients with higher severity. Baseline pain intensity demonstrated a clear stepwise gradient across severity strata. Mean lowest, average, and highest 24- hour pain intensity values increased progressively from CPGS I/II to CPGS IV. A similar gradient was observed for pain-related disability measured with the modified Pain Disability Index (mPDI), with mean baseline scores increasing from CPGS I/II to CPGS IV and the proportion of patients with severe disability (mPDI ≥50) rising accordingly. Neuropathic pain characteristics measured by PDQ-7 were more prevalent with increasing severity, and sleep disturbances related to spinal pain were also reported more frequently in higher CPGS strata. Finally, comorbidity burden and comedication use were broadly comparable across groups, although the prevalence of multiple comorbidities was slightly higher in patients with CPGS III and CPGS IV.

Patient-Reported Health Status, Psychological Burden, and Quality of Life Baseline patient-reported outcomes across CPGS severity strata are summarized in Table 2. Clear gradients across severity groups were observed for functional impairment, psychosocial burden, and health-related quality of life. Functional capacity assessed with the Hannover Functional Ability Questionnaire (HFAQ) decreased progressively with increasing CPGS severity. Mean HFAQ scores declined from 68.5 (SD 23.0) in CPGS I/II to 58.4 (22.7) in CPGS III and 50.7 (22.4) in CPGS IV, with the proportion of patients showing marked functional impairment (HFAQ <65) increasing accordingly. Fear-avoidance beliefs showed a similar pattern. Both physical activity (PA) –related fear avoidance (FABQ-PA) and work (W)-related fear avoidance (FABQ-W) increased with higher CPGS severity, with a growing proportion of patients exceeding clinically relevant thresholds. Pain-related self-efficacy assessed by the Pain Self-Efficacy Questionnaire (PSEQ) decreased markedly across severity strata. Mean PSEQ scores declined from 34.7 (SD 15.1) in CPGS I/II patients to 29.9 (12.7) in CPGS III and 26.6 (11.6) in CPGS IV, with reduced self-efficacy (PSEQ ≤30) becoming progressively more frequent. General well-being measured by the MQHHF and health-related quality of life assessed with the VR- 12 also deteriorated with increasing severity. Physical health status (VR-12 PCS) showed substantial impairment across all groups and declined further with increasing CPGS severity, while mental health status (VR-12 MCS) also decreased, although less steeply. Psychological distress assessed with DASS-21 increased consistently across severity strata. Mean depression, anxiety, and stress scores were highest in the CPGS IV group, and the proportion of patients exceeding thresholds for strong to extremely severe symptoms increased accordingly.

Previous Interventional and Pharmacological Treatment Exposure

Previous interventional procedures and pharmacological treatment exposure prior to enrollment in the outpatient multimodal treatment program are summarized in Table 3. Across all CPGS severity strata, patients had been exposed to extensive prior treatment. Interventional procedures were common and increased with higher CPGS severity. Neural therapy had been performed in 65.1% of patients in the CPGS I/II group, 74.1% in CPGS III, and 82.0% in CPGS IV, while previous periradicular therapy was documented in 32.0%, 40.6%, and 49.9% of patients, respectively. Prior spine surgery related to low back pain was reported in 31.1% of patients in CPGS I/II, increasing to 37.6% in CPGS III and 42.7% in CPGS IV. All patients had received pharmacological pain treatment prior to study entry. Non-steroidal anti-inflammatory drugs were used by all patients, and the majority had been treated with multiple additional medication classes, including non-opioid analgesics, opioids, antidepressants, muscle relaxants, and antiepileptic drugs. Use of mild opioids was reported by approximately 78–82% of patients, while strong opioids had been prescribed in approximately 43–46% of cases across severity strata. Overall pharmacological treatment exposure was substantial, with patients receiving a mean of approximately nine different background medications prior to study entry. Rescue medication use was also common and reported by roughly two-thirds of patients across all CPGS groups.

Current Pharmacological Treatment at Baseline

Baseline pharmacological treatment at program entry is summarized in Table 4. Overall, patients were receiving substantial ongoing pharmacological treatment at the time of program entry. The mean number of concurrent background medications was comparable across severity strata, averaging approximately 4.5 drugs per patient. Non-steroidal anti-inflammatory drugs (NSAIDs) represented the most frequently used medication class and were reported by roughly three quarters of patients across all CPGS groups. Non-opioid analgesics were used by approximately 40% of patients. Opioid therapy was common. Mild opioids were reported in approximately 45–47% of patients, whereas strong opioids were used by roughly one quarter of the cohort across all severity strata. Cannabinoid use was comparatively uncommon and reported by about 9% of patients.

Adjuvant analgesic medications were widely prescribed. Muscle relaxants were used by approximately three quarters of patients, antidepressants by around 70%, and antiepileptic drugs by roughly one quarter of the cohort. Approximately half of patients across all severity groups reported the current use of rescue medications.

Primary Endpoint: Change in Pain-Related Disability (mPDI)

Treatment-associated changes in pain-related disability are summarized in Table 5 and Figures 2-4. At baseline, mean mPDI scores differed substantially across CPGS severity strata, increasing from 42.9 (SD 16.9; 95% CI 42.2–43.6) in CPGS I/II patients to 56.0 (15.4; 95% CI 55.4–56.6) in CPGS III and 65.1 (14.9; 95% CI 64.6–65.5) in CPGS IV (between-group comparison $p < 0.001$; partial $\eta^2 = 0.253$). After three weeks of interdisciplinary outpatient multimodal treatment, mean mPDI scores decreased markedly across all severity strata, reaching 24.8 (SD 20.2; 95% CI 24.0–25.6) in CPGS I/II patients, 26.0 (20.0; 95% CI 25.2–26.8) in CPGS III, and 28.1 (20.9; 95% CI 27.5–28.7) in CPGS IV patients. Within-group reductions were statistically significant in all groups (all $p < 0.001$) and associated with effect sizes increasing with baseline severity (Cohen's $d = 0.974$ for CPGS I/II, 1.681 for CPGS III, and 2.037 for CPGS IV). Mean absolute improvements in mPDI increased with baseline severity, amounting to 18.1 mm (SD 24.9; 95% CI 17.1–19.1) in CPGS I/II patients, 30.0 mm (24.7; 95% CI 29.1–31.0) in CPGS III, and 37.0 mm (24.8; 95% CI 36.2–37.7) in CPGS IV (between-group comparison $p < 0.001$; partial $\eta^2 = 0.088$). Mean relative improvements were 42.3% (SD 44.8), 51.4% (38.6), and 54.9% (34.9), respectively (between-group comparison $p < 0.001$; partial $\eta^2 = 0.018$). Responder analyses confirmed clinically meaningful improvements across all severity strata. Using the predefined absolute improvement criterion of ≥20 mm in mPDI, 48.5% of patients in the CPGS I/II group (n = 1,175), 66.8% in the CPGS III group (n = 1,723), and 74.1% in the CPGS IV group (n = 3,124) met the responder definition (between-group comparison $p < 0.001$; $\varphi = 0.222$). Applying the relative improvement criterion of ≥50% reduction in mPDI, responder rates were 50.1% (n = 1,214) in CPGS I/II patients, 56.7% (n = 1,462) in CPGS III patients, and 59.3% (n = 2,497) in CPGS IV patients (between-group comparison $p < 0.001$; $\varphi = 0.075$). Consistent with these findings, the proportion of patients with severe disability (mPDI ≥50) decreased markedly from baseline to week 3 across all severity strata (Table 5 and Figures 2-4).

Multivariate Covariance Analyses of Treatment-Associated Change Profiles Across CPGS Severity Strata

Multivariate analyses of covariance (MANCOVA) were conducted to examine whether treatment-associated change profiles across multiple outcome domains differed between CPGS severity strata while adjusting for treatment center, age, and sex (Table 6). In the multidimensional outcome model (Table 6a), which simultaneously included change scores for disability, pain intensity, functional capacity, health-related quality of life, and psychological distress domains, the overall multivariate effect of CPGS group was statistically significant (Wilks' $\Lambda = 0.880$, $F(14, 18{,}410) = 86.41$, $p < 0.001$; partial $\eta^2 = 0.062$). Comparable results were obtained using Pillai's trace, indicating a robust association between baseline severity group and the overall pattern of treatment-associated changes across outcome domains. Age and sex showed statistically significant but small multivariate effects (partial $\eta^2 = 0.027$ and 0.013, respectively), whereas treatment center was not significantly associated with the multivariate change pattern ($p = 0.762$), indicating that treatment-associated improvements were consistent across participating centers. A second MANCOVA model based on a predefined clinical core outcome set (Table 6b), including change scores for average pain intensity and pain-related disability, yielded convergent findings. The multivariate effect of CPGS group remained statistically significant (Wilks' $\Lambda = 0.909$, $F(4, 18{,}420) = 224.13$, $p < 0.001$; partial $\eta^2 = 0.046$). Age and sex again demonstrated small but statistically significant multivariate effects (partial $\eta^2 = 0.008$ each), whereas treatment center was not significantly associated with treatment-related change patterns ($p = 0.334$) (Table 6a and Table 6b).

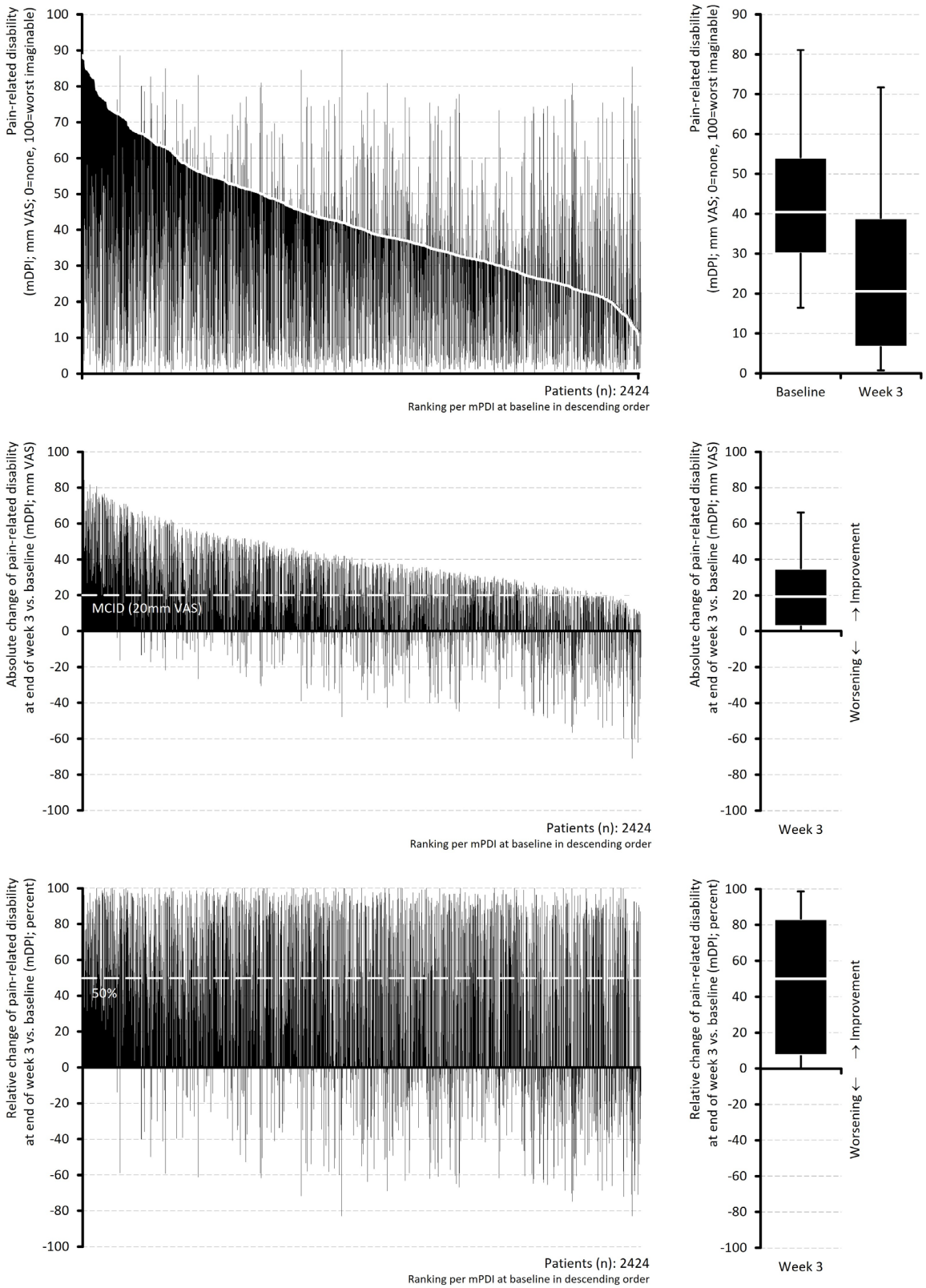

Figure 2: Treatment effects on pain-related disability (mPDI) for patients classified as CPGS I/II.
Upper panel – left side: change of the modified pain disability index (mPDI; mm VAS) from baseline (white line; patients ranked from left to right by descending disability at baseline) to end of week 3 (black columns); right side: corresponding box-and-whisker plot (box: 25,50, 75% percentiles; whiskers: 2.5 and 97.5% percentiles) for mPDI at baseline and end of week 24. Middle panel – left: absolute mPDI improvements (above zero-line) and worsening's (below zero-line) until end of week 3 vs. baseline (incl. dashed white line for the minimal clinical important difference threshold of 20mm VAS); right: corresponding box-and-whisker plot. Lower panel – left: relative (percent) mPDI improvements (above zero-line) and worsening's (below zero-line) until end of week 3 vs. baseline (incl. dashed white line for the 50% improvement); right: corresponding box-and-whisker plot.

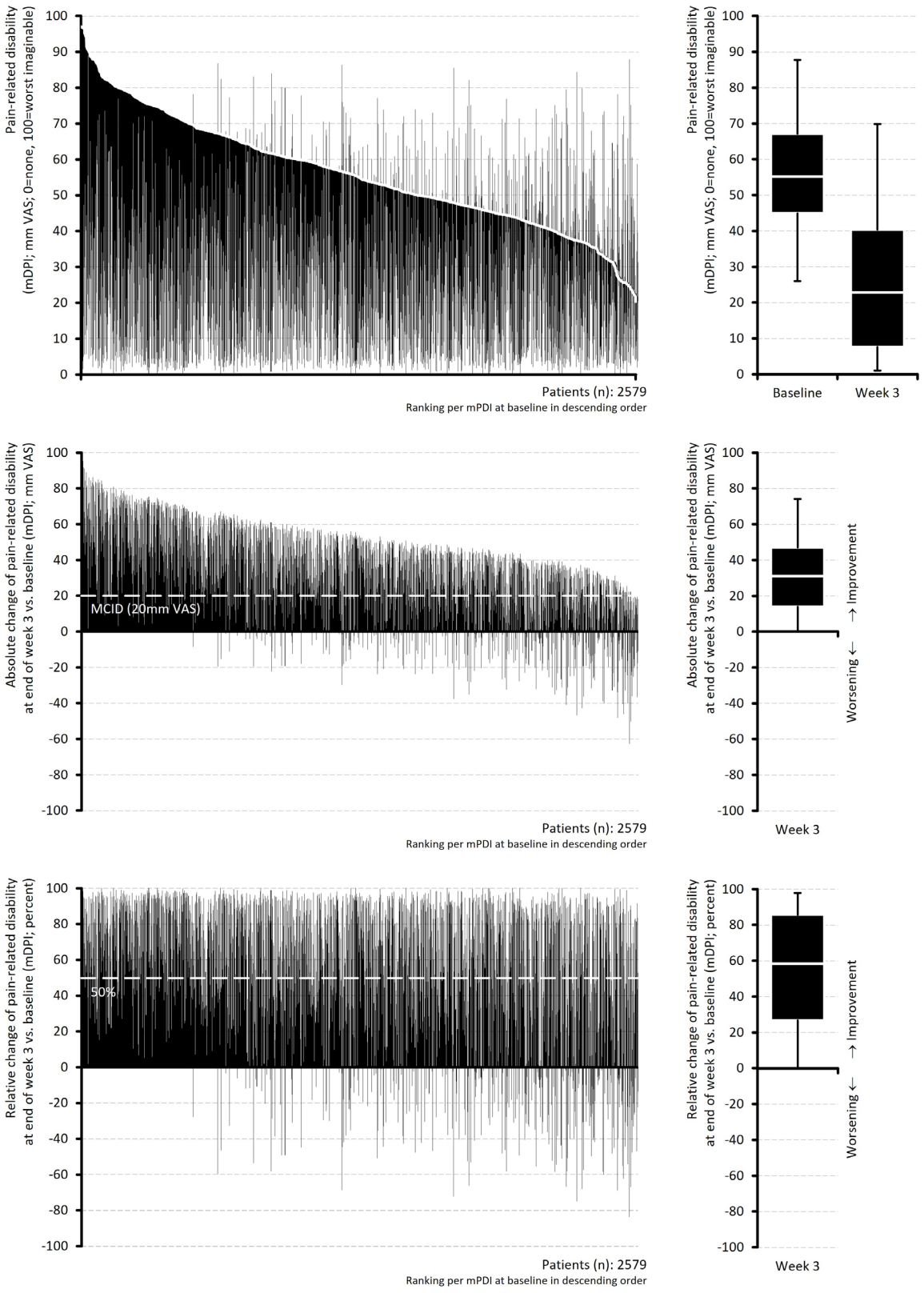

Figure 3: Treatment effects on pain-related disability (mPDI) for patients classified as CPGS III.
Upper panel – left side: change of the modified pain disability index (mPDI; mm VAS) from baseline (white line; patients ranked from left to right by descending disability at baseline) to end of week 3 (black columns); right side: corresponding box-and-whisker plot (box: 25,50, 75% percentiles; whiskers: 2.5 and 97.5% percentiles) for mPDI at baseline and end of week 24. Middle panel – left: absolute mPDI improvements (above zero-line) and worsening's (below zero-line) until end of week 3 vs. baseline (incl. dashed white line for the minimal clinical important difference threshold of 20mm VAS); right: corresponding box-and-whisker plot. Lower panel – left: relative (percent) mPDI improvements (above zero-line) and worsening's (below zero-line) until end of week 3 vs. baseline (incl. dashed white line for the 50% improvement); right: corresponding box-and-whisker plot.

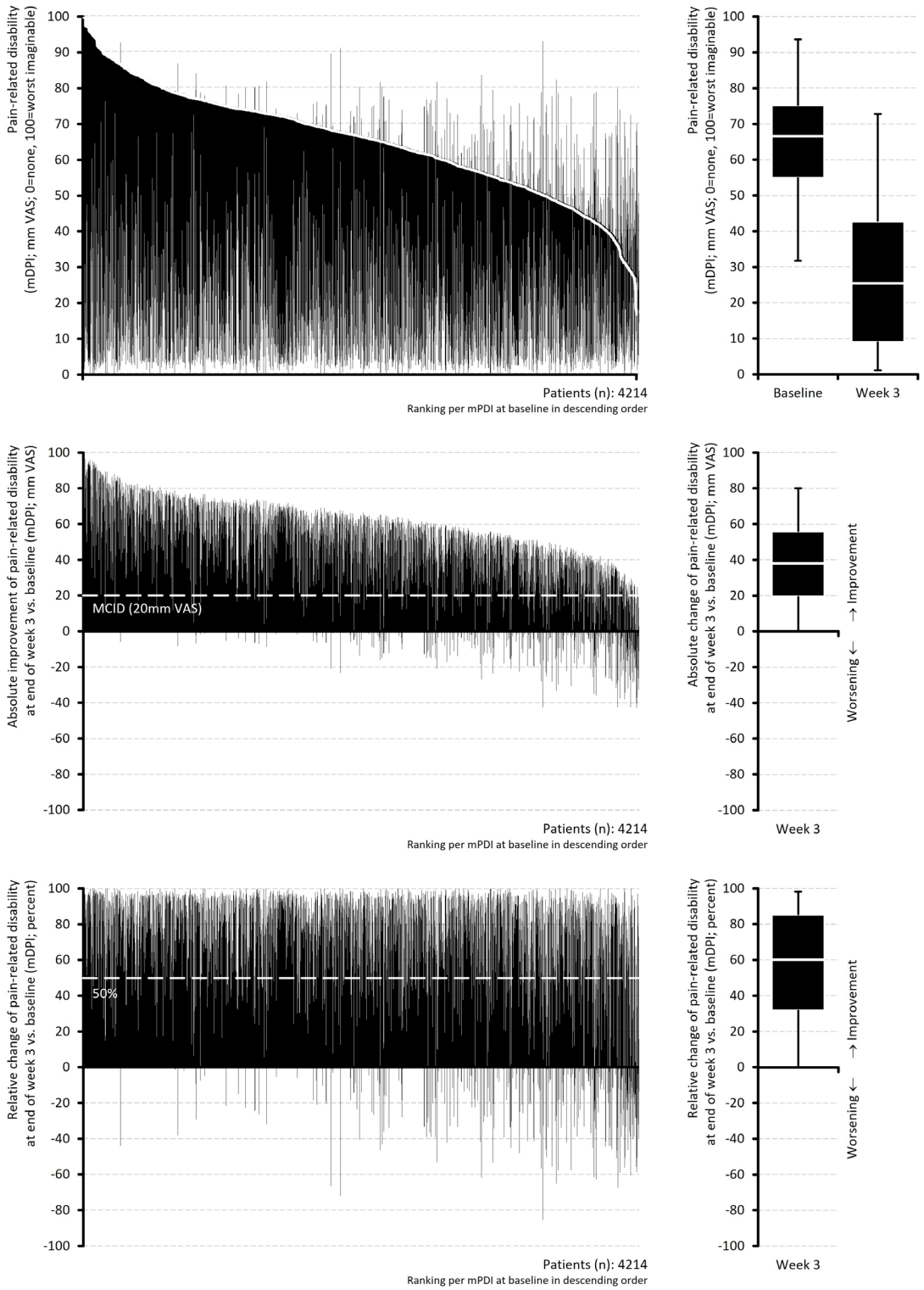

Figure 4: Treatment effects on pain-related disability (mPDI) for patients classified as CPGS IV.
Upper panel – left side: change of the modified pain disability index (mPDI; mm VAS) from baseline (white line; patients ranked from left to right by descending disability at baseline) to end of week 3 (black columns); right side: corresponding box-and-whisker plot (box: 25,50, 75% percentiles; whiskers: 2.5 and 97.5% percentiles) for mPDI at baseline and end of week 24. Middle panel – left: absolute mPDI improvements (above zero-line) and worsening's (below zero-line) until end of week 3 vs. baseline (incl. dashed white line for the minimal clinical important difference threshold of 20mm VAS); right: corresponding box-and-whisker plot. Lower panel – left: relative (percent) mPDI improvements (above zero-line) and worsening's (below zero-line) until end of week 3 vs. baseline (incl. dashed white line for the 50% improvement); right: corresponding box-and-whisker plot.

Secondary endpoints

Treatment-associated changes across secondary outcome domains are summarized in Tables 7–11 and Figure 5 (effect sizes). These analyses included measures of pain intensity, general well-being, health-related quality of life, psychological distress, functional capacity, fear-avoidance beliefs, pain self-efficacy, and pharmacological treatment exposure. Across all outcome domains, statistically significant improvements were observed between baseline (T0) and week 3 (T3) in all predefined CPGS severity strata. While baseline values differed substantially between severity groups, the direction of treatment-associated changes was consistent across strata (Tables 7–11 and Figure 5).

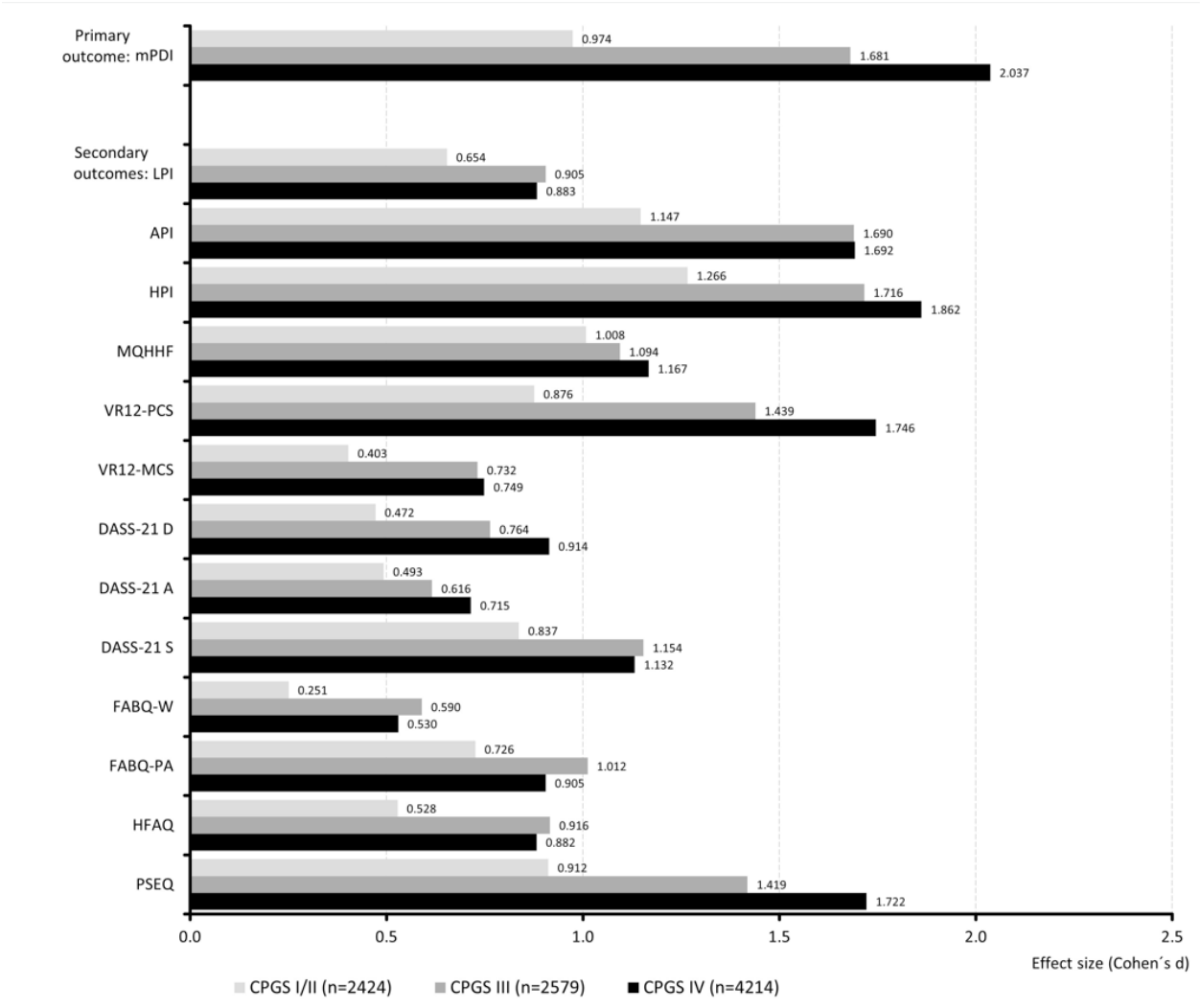

Figure 5: Standardized treatment-associated changes across primary and secondary outcome domains after completion of the 3-week interdisciplinary outpatient multimodal treatment program. Effect sizes are shown as Cohen's d (standardized mean difference, SMD) for within-group pre–post changes (baseline – end of week 3) stratified by baseline CPGS severity group. Positive values indicate improvement across all domains after directionality adjustment of the respective scales.

Table 5: Primary endpoint: changes in pain-related disability (mPDI) from baseline to week three.

Cohort	CPGS I/II	CPGS III	CPGS IV	for all	I/II vs. III	I/II vs. IV	III vs. IV
Number of patients [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	-	-	-	-
mPDI at baseline [mm VAS; mean (SD);95% CI]	42.9 (16.9)	56.0 (15.4)	65.1 (14.9)	<0.001	<0.001	<0.001	<0.001
	42.2-43.6	55.4-56.6	64.6-65.5	-0.253	-0.143	-0.317	-0.078
mPDI at end of week three [mm VAS; mean (SD);95% C)]	24.8 (20.2)	26.0 (20.0)	28.1 (20.9)	<0.001	0.111	<0.001	<0.001
	24.0-25.6	25.2-26.8	27.5-28.7	-0.005	-0.001	-0.006	-0.003

Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001				
	-0.974	-1.681	-2.037				
Absolute mPDI improvement vs. baseline [mm VAS; mean (SD);95% CI]	18.1 (24.9)	30.0 (24.7)	37.0 (24.8)	<0.001	<0.001	<0.001	<0.001
	17.1-19.1	29.1-31.0	36.2-37.7	-0.088	-0.48	-0.761	-0.283
Relative mPDI improvement vs. baseline [%; mean (SD);95% CI]	42.3 (44.8)	51.4 (38.6)	54.9 (34.9)	<0.001	<0.001	<0.001	0.001
	40.5-44.1	50.0-52.9	53.9-56.0	-0.018	-0.218	-0.325	-0.096

A: Continuous outcome analysis

Baseline (T0) and week-three (T3) pain-related disability (mPDI) scores and absolute change values across CPGS severity groups. Continuous variables are presented as mean (SD) with 95% confidence intervals. Between-group comparisons were performed using analysis of covariance (ANCOVA) adjusting for center. Effect sizes are reported as partial eta-squared (η^2) for continuous between-group comparisons (ANCOVA) and Cohen´s d for within-group pre-post changes.

Figure 5:

Cohort	CPGS I/II	CPGS III	CPGS IV	for all	I/II vs. III	I/II vs. IV	III vs. IV
Number of patients [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	-	-	-	-
mPDI ≥ 50 mm VAS at baseline [n (%)]	790	1578	3514	<0.001	<0.001	<0.001	<0.001
	-32.6	-61.2	-83.4	-0.432	-0.286	-0.512	-0.249
mPDI ≥ 50 mm VAS at end of week three [n (%)]	305	350	691	<0.001	0.3	<0.001	0.002
	-12.6	-13.6	-16.4	-0.048	-0.029	-0.103	-0.076
Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001				
	-0.667	-1.271	-1.788				
Absolute mPDI improvement vs. baseline ≥20 mm VAS [n (%)]	1175 (48.5)	1723 (66.8)	3124 (74.1)	<0.001	<0.001	<0.001	<0.001
				-0.222	0.186	0.259	0.079
Relative mPDI improvement vs baseline ≥50 percent [n (%)]	1214 (50.1)	1462 (56.7)	2497 (59.3)	<0.001	<0.001	<0.001	0.037
				-0.075	0.066	0.089	0.025

B: Responder analysis

Responder rates based on predefined improvement thresholds in mPDI across CPGS severity groups. The primary responder definition was ≥20-point absolute improvement and/or ≥30% relative improvement from baseline. Additional categorical thresholds are presented where applicable. Between-group comparisons were conducted using adjusted chi-square analyses. Effect sizes are reported as phi coefficients or Cramer's V for categorical between-group comparisons, and Cohen´s d for within-group pre-post changes.

Table 6: Multivariate analyses of covariance (MANCOVA) examining treatment-associated changes across CPGS severity groups.

Effect	Multivariate test	Value	F	Hypothesis df	Error df	p-value	Partial η^2
Center	Wilks' Λ	1	0.593	7	9205	0.762	0
	Pillai´s Trace	0	0.559	7	9207	0.789	0
Age	Wilks' Λ	0.973	35.966	7	9205	<0.001	0.027
	Pillai´s Trace	0.027	35.966	7	9205	<0.001	0.027
Sex	Wilks' Λ	0.987	16.92	7	9205	<0.001	0.013
	Pillai´s Trace	0.013	16.92	7	9205	<0.001	0.013
CPGS group	Wilks' Λ	0.88	86.573	14	18414	<0.001	0.062
	Pillai´s Trace	0.12	84.325	14	18416	<0.001	0.06

6a. Multidimensional outcome model

(ΔmPDI, ΔAPI, ΔHFAQ, ΔVR12-PCS, ΔVR12-MCS, ΔDASS-D, ΔPSEQ).

Table 6b: Clinical core outcome model (ΔAPI, ΔmPDI).

Effect	Multivariate test	Value	F	Hypothesis df	Error df	p-value	Partial η^2
Center	Wilks' Λ	1	1.096	2	9210	0.334	0
	Pillai´s Trace	0	1.096	2	9210	0.334	0

Age	Wilks' Λ	0.992	37.924	2	9210	<0.001	0.008
	Pillai´s Trace	0.008	37.924	2	9210	<0.001	0.008
Sex	Wilks' Λ	0.992	35.687	2	9210	<0.001	0.008
	Pillai´s Trace	0.008	35.687	2	9210	<0.001	0.008
CPGS group	Wilks' Λ	0.909	224.126	4	18420	<0.001	0.046
	Pillai´s Trace	0.091	219.255	4	18422	<0.001	0.045

Multivariate analyses of covariance (MANCOVA) evaluating differences in treatment-associated change scores (T3–T0) across CPGS severity groups. Table 6a presents a multidimensional model including changes (delta, D) in pain-related disability (mPDI), average pain intensity (API), functional capacity (HFAQ), physical and mental health-related quality of life (VR12-PCS, VR12-MCS), depressive symptoms (DASS-21-D), and pain self-efficacy (PSEQ). Table 6b presents a focused clinical core model including combined changes in API and mPDI only. All models were adjusted for treatment center, age, and sex. Multivariate test statistics are reported as Wilks' Lambda and Pillai's Trace. Effect sizes are reported as partial eta-squared (η^2).

Table 7: Changes in pain intensity measures from baseline to week three.

	CPGS I/II	CPGS III	CPGS IV	Between cohort significance (effect size: partial eta-squared / Cramer V)
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	-
Lowest 24-hr. pain intensity at baseline [LPI; mm VAS; mean (SD)]	17.9 (13.8)	23.1 (16.5)	26.2 (19.6)	<0.001 (0.037)
Lowest 24-hr. pain intensity at end of week three [LPI; mm VAS; mean (SD)]	8.9 (13.8)	9.2 (14.1)	10.5 (15.8)	<0.001 (0.003)
Absolute LPI improvement vs. baseline [mm VAS; mean (SD)]	9.0 (18.5)	13.9 (20.3)	15.7 (22.9)	<0.001 (0.017)
Relative LPI improvement vs. baseline [%; mean (SD)]	53.2 (47.8)	59.5 (45.0)	57.8 (46.3)	<0.001 (0.003)
Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001	
	-0.654	-0.905	-0.883	
LPI ≥ 30 mm VAS at ...baseline [n (%)]	407 (16.8)	720 (27.9)	1471 (34.9)	<0.001 (0.165)
...end of week three [n (%)]	235 (9.7)	240 (9.3)	516 (12.2)	<0.001 (0.044)
Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001	
	-0.21	-0.492	-0.554	
Absolute LPI improvement vs. baseline ≥20 mm VAS [n (%)]	636 (26.2)	899 (34.9)	1646 (39.1)	<0.001 (0.110)
Relative LPI improvement vs. baseline ≥50 percent [n (%)]	1511 (62.3)	1745 (67.7)	2774 (65.8)	<0.001 (0.042)
Average 24-hr. pain intensity at baseline [API; mm VAS; mean (SD)]	42.5 (19.0)	51.8 (18.2)	55.2 (19.3)	<0.001 (0.071)
Average 24-hr. pain intensity at end of week three [API; mm VAS; mean (SD)]	20.9 (18.6)	20.8 (18.5)	22.5 (19.5)	<0.001 (0.002)
Absolute API improvement vs. baseline [mm VAS; mean (SD)]	21.6 (25.1)	31.0 (24.8)	32.7 (25.8)	<0.001 (0.033)
Relative API improvement vs. baseline [%; mean (SD)]	48.9 (42.7)	57.2 (38.3)	56.5 (37.9)	<0.001 (0.008)
Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001	
	-1.147	-1.69	-1.692	
API ≥ 50 mm VAS at ...baseline [n (%)]	964 (39.8)	1520 (58.9)	2724 (64.6)	<0.001 (0.207)
...end of week three [n (%)] Within cohort significance (effect size; Cohen´s d)	218 (9.0)	247 (9.6)	437 (10.4)	0.175 (0.019)
	<0.001	<0.001	<0.001	
	-0.768	-1.218	-1.354	
Absolute API improvement vs. baseline ≥20 mm VAS [n (%)]	1303 (53.8)	1801 (69.8)	2914 (69.2)	<0.001 (0.145)
Relative API improvement vs. baseline ≥50 percent [n (%)]	1388 (57.3)	1678 (65.1)	2655 (63.0)	<0.001 (0.062)

Highest 24-hr. pain intensity at baseline [HPI; mm VAS; mean (SD)]	66.2 (21.0)	75.0 (18.9)	79.0 (17.1)	<0.001 (0.072)
Highest 24-hr. pain intensity at end of week three [HPI; mm VAS; mean (SD)]	37.2 (24.8)	37.3 (24.7)	38.6 (25.4)	0.026 (0.001)
Absolute HPI improvement vs. baseline [mm VAS; mean (SD)]	29.1 (31.2)	37.8 (30.2)	40.3 (29.5)	<0.001 (0.023)
Relative HPI improvement vs. baseline [%; mean (SD)]	40.0 (43.7)	47.3 (38.5)	48.7 (36.4)	<0.001 (0.009)
Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001	
	-1.266	-1.716	-1.862	
HPI ≥ 80 mm VAS at ...baseline [n (%)]	784 (32.3)	1325 (51.4)	2537 (60.2)	<0.001 (0.228)
...end of week three [n (%)] Within cohort significance (effect size; Cohen´s d)	91 (3.8)	101 (3.9)	214 (5.1)	0.015 (0.030)
	<0.001	<0.001	<0.001	
	-0.801	-1.252	-1.453	
Absolute API improvement vs. baseline ≥20 mm VAS [n (%)]	1525 (62.9)	1864 (72.3)	3086 (73.2)	<0.001 (0.096)
Relative API improvement vs. baseline ≥50 percent [n (%)]	1124 (46.4)	1320 (51.2)	2143 (50.9)	<0.001 (0.041)

Changes in lowest (LPI), average (API), and highest (HPI) 24-hour pain intensity scores from baseline (T0) to end of week three (T3) across CPGS severity groups. Absolute and relative change scores are reported as mean (SD); categorical thresholds as n (%). Between-group comparisons were performed using ANCOVA adjusting for center. Partial eta-squared is reported for continuous between-group comparisons (ANCOVA), Cramer's V for categorical between-group comparisons, and Cohen's d for within-group pre-post changes. SD = standard deviation; ANCOVA = analysis of covariance.

Table 8: Changes in general well-being and health-related quality of life.

	CPGS I/II	CPGS III	CPGS IV	Between cohort significance (effect size: partial eta-squared / Cramer V)
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	-
Generall wellbeing at ...baseline (MQHHF; NRS-35; mean (SD)]	15.2 (7.1)	11.6 (7.0)	9.5 (6.4)	<0.001 (0.106)
end of week three [MQHHF; NRS-35; mean (SD)]	22.3 (6.9)	19.7 (7.7)	17.8 (7.8)	<0.001 (0.055)
Absolute improvement VS. baseline [NRS-35; mean (SD)]	7.1 (7.6)	8.1 (7.9)	8.3 (8.2)	<0.001 (0.004)
Relative Improvement VS. baseline [%; mean (SD)]	33.8 (33.2)	31.8 (34.0)	29.8 (34.2)	<0.001 (0.002)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(1.008)	(1.094)	(1.167)	
MQHHF <10 at baseline [n (%)]	634 (26.2)	1235 (47.9)	2498 (59.3)	<0.001 (0.271)
end of week three [n (%)]	143 (5.9)	343 (13.3)	800 (19.0)	<0.001 (0.155)
Within cohort significance (effect size; Cohen's d)	<0.001	<0.001	<0.001	
	(0.574)	(0.810)	(0.906)	
Physical quality-of-life at .baseline [VR12-PCS; mean (SD)]	35.3 (9.5)	30.7 (8.6)	27.4 (8.3)	<0.001 (0.121)
...end of week three [VR12-PCS; mean (SD)]	43.8 (10.0)	44.0 (9.9)	43.4 (10.0)	<0.001 (0.001)
Absolute improvement VS. baseline [mean (SD)]	8.5 (13.7)	13.3 (12.7)	16.0 (12.9)	<0.001 (0.052)
Relative Improvement VS. baseline [%; mean (SD)]	38.6 (69.6)	59.5 (78.6)	76.8 (80.9)	<0.001 (0.039)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.876)	(1.439)	(1.746)	
VR12-PCS<35 at baseline [n (%)]	1051 (43.4)	1785 (69.2)	3475 (82.5)	<0.001 (0.344)

...end of week three [n (%)]	455 (18.8)	479 (18.6)	883 (21.0)	0.029 (0.029)
Within cohort significance(effect size; Cohen's d)	<0.001	<0.001	<0.001	
	(0.551)	(1.186)	(1.562)	
Mental quality-of-life at ...baseline [VR12-MCS; mean (SD)]	43.2 (11.8)	38.7 (11.4)	38.0 (11.5)	<0.001 (0.035)
...end of week three [VR12-MCS; mean (SD)]	47.6 (10.1)	46.6 (10.4)	46.3 (10.6)	<0.001 (0.003)
Absolute improvement vs. baseline [mm VAS; mean (SD)]	4.4 (14.7)	8.0 (14.9)	8.3 (15.2)	<0.001 (0.012)
Relative Improvement vs. baseline [%; mean (SD)]	19.8 (46.4)	32.4 (55.3)	35.3 (62.8)	<0.001 (0.013)
Within cohort significance (effect size; Cohen's d)	<0.001	<0.001	<0.001	
	(0.403)	(0.732)	(0.749)	
VR12-MCS≤40 at ...baseline [n (%)]	985 (40.6)	1527 (59.2)	2524 (59.9)	<0.001 (0.168)
...end of week three [n (%)]	546 (22.5)	664 (25.7)	1156 (27.4)	<0.001 (0.046)
Within cohort significance (effect size; Cohen´s d)	<0.001	<0.001	<0.001	
	-0.397	-0.719	-0.693	

Changes in general well-being (MQHHF), physical health-related quality of life (VR12-PCS), and mental health-related quality of life (VR12-MCS) from baseline to week three. Continuous variables are presented as mean (SD); categorical impairment thresholds as n (%). Analyses were adjusted for center using ANCOVA. Partial eta-squared is reported for continuous between-group comparisons (ANCOVA), Cramer's V for categorical between-group comparisons, and Cohen's d for within-group pre-post changes.

Table 9: Changes in psychological distress.

	CPGS I/II	CPGS III	CPGS IV	Between cohort significance (effect size: partial eta-squared / Cramer V)
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	
Depression at ...baseline [DASS-21 D; NRS-21; mean (SD)]	5.1 (4.9)	6.7 (5.2)	7.4 (5.2)	<0.001 (0.034)
...end of week three [DASS-21 D; NRS-21; mean (SD)]	3.1 (3.2)	3.4 (3.4)	3.4 (3.3)	0.005 (0.001)
Absolute improvement vs. baseline [NRS-21; mean (SD)]	1.9 (5.6)	3.3 (5.9)	4.0 (5.9)	<0.001 (0.021)
Relative Improvement vs. baseline [%; mean (SD)]	33.7 (47.1)	42.2 (46.1)	45.8 (44.5)	<0.001 (0.012)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.472)	(0.764)	(0.914)	
Strong to extreme depression (DASS-21 D≥11) at ...baseline [n (%)]	331 (13.7)	568 (22.0)	1133 (26.9)	<0.001 (0.130)
...end of week three [n (%)]	107 (4.4)	133 (5.2)	213 (5.1)	0.407 (0.014)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.327)	(0.508)	(0.624)	
Anxiety at ...baseline [DASS-21 A; NRS-21; mean (SD)]	3.7 (4.6)	4.3 (4.9)	4.6 (4.7)	<0.001 (0.006)
...end of week three [DASS-21 A; NRS-21; mean (SD)]	2.0 (2.0)	2.0 (2.0)	2.0 (2.1)	0.450 (<0.001)
Absolute improvement vs. baseline [NRS-21; mean (SD)]	1.8 (4.9)	2.3 (5.2)	2.6 (4.9)	<0.001 (0.005)
Relative Improvement vs. baseline [%; mean (SD)]	32.8 (45.4)	38.6 (46.6)	41.5 (45.2)	<0.001 (0.006)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.493)	(0.616)	(0.715)	

Strong to extreme anxiety (DASS-21 A≥8) at ...baseline [n (%)]	370 (15.3)	510 (19.8)	1006 (23.9)	<0.001 (0.088)
...end of week three [n (%)]	57 (2.4)	59 (2.3)	125 (3.0)	0.150 (0.020)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.468)	(0.581)	(0.644)	
Stress at ...baseline [DASS-21 S; NRS-21; mean (SD)]	7.3 (5.1)	9.1 (5.3)	9.1 (5.2)	<0.001 (0.023)
...end of week three [DASS-21 S; NRS-21; mean (SD)]	3.7 (3.4)	3.9 (3.5)	4.1 (3.6)	<0.001 (0.002)
Absolute improvement vs. baseline [NRS-21; mean (SD)]	3.6 (5.8)	5.2 (6.2)	5.0 (6.0)	<0.001 (0.012)
Relative Improvement vs. baseline [%; mean (SD)]	41.6 (43.8)	48.9 (42.6)	48.0 (42.2)	<0.001 (0.005)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.837)	(1.154)	(1.132)	
Strong to extreme stress (DASS-21 S≥13) at ...baseline [n (%)]	444 (18.3)	639 (24.8)	1146 (27.2)	<0.001 (0.085)
...end of week three [n (%)]	76 (3.1)	81 (3.1)	165 (3.9)	0.129 (0.021)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen's d)	(0.506)	(0.657)	(0.678)	

Changes in depressive symptoms (DASS-21-D), anxiety (DASS-21-A), and stress (DASS-21-S) from baseline to week three. Data are presented as mean (SD) or n (%). Between-group differences were evaluated using ANCOVA with center as covariate. Partial eta-squared is reported for continuous between-group comparisons (ANCOVA), Cramer's V for categorical between-group comparisons, and Cohen's d for within-group pre-post changes.

Table 10: Changes in fear-avoidance beliefs, functional capacity, and self-efficacy.

	CPGS I/II	CPGS III	CPGS IV	Between cohort significance (effect size: partial eta-squared / Cramer V)
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	-
FABQ - work at ...baseline [NRS-42; mean (SD)]	13.8 (11.6)	18.3 (12.4)	20.4 (11.8)	<0.001 (0.049)
...end of week three [NRS-42; mean (SD)]	11.2 (8.7)	12.0 (8.9)	14.5 (10.6)	<0.001 (0.022)
Absolute improvement vs. baseline [NRS-42; mean (SD)]	2.6 (12.2)	6.4 (12.5)	5.9 (11.8)	<0.001 (0.016)
Relative Improvement vs. baseline [%; mean (SD)]	24.5 (38.4)	33.4 (38.4)	32.1 (37.0)	<0.001 (0.009)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.251	-0.59	-0.53	
FABQ-W≥34 at ...baseline [n (%)]	184 (7.6)	368 (14.3)	690 (16.4)	<0.001 (0.106)
...end of week three [n (%)]	24 (1.0)	40 (1.6)	212 (5.0)	<0.001 (0.118)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.33	-0.485	-0.373	
FABQ - physical activity at ...baseline (NRS-24; mean (SD)]	13.1 (6.0)	15.2 (9.3)	15.8 (5.5)	<0.001 (0.035)
...end of week three [NRS-24; mean (SD)]	8.7 (5.9)	9.3 (5.7)	10.5 (6.1)	<0.001 (0.015)
Absolute improvement vs. baseline [NRS-24; mean (SD)]	4.3 (7.2)	5.9 (7.4)	5.3 (7.7)	<0.001 (0.006)
Relative Improvement vs. baseline [%; mean (SD)]	36.4 (36.1)	39.3 (34.2)	35.8 (34.3)	<0.001 (0.002)
Within cohort significance	<0.001	<0.001	<0.001	

(effect size; Cohen´s d)	-0.726	-1.012	-0.905	
FABQ-PA≥15 at …baseline [n (%)]	1088 (44.9)	1475 (57.2)	2689 (63.8)	<0.001 (0.156)
…end of week three [n (%)]	360 (14.9)	373 (14.5)	938 (22.3)	<0.001 (0.087)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.695	-0.995	-0.925	
HFQB at … baseline [NRS-100; mean (SD)]	68.5 (23.0)	58.4 (22.7)	50.7 (22.4)	<0.001 (0.094)
...end of week three [HFQB; NRS-100; mean (SD)]	79.1 (16.5)	77.0 (17.6)	69.9 (21.1)	<0.001 (0.044)
Absolute improvement vs. baseline [NRS-100; mean (SD)]	10.6 (22.4)	18.6 (23.2)	19.2 (24.5)	<0.001 (0.024)
Relative Improvement vs. baseline [%; mean (SD)]	30.9 (39.0)	39.5 (37.4)	35.8 (37.9)	<0.001 (0.007)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.528	-0.916	-0.882	
HFQB<65 at …baseline [n (%)]	815 (33.6)	1512 (58.6)	3066 (72.8)	<0.001 (0.325)
…end of week three [n (%)]	394 (16.3)	550 (21.3)	1337 (31.7)	<0.001 (0.155)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.41	-0.824	-0.901	
Pain self efficacy at …baseline [PSEQ; NRS-60; mean (SD)]	34.7 (15.1)	29.9 (12.7)	26.6 (11.6)	<0.001 (0.062)
...end of week three [PSEQ; NRS-60; mean (SD)]	46.4 (10.1)	46.2 (10.1)	45.6 (10.5)	0.015 (0.001)
Absolute improvement vs. baseline [NRS-60; mean (SD)]	11.7 (17.8)	16.3 (15.8)	19.1 (15.1)	<0.001 (0.034)
Relative Improvement vs. baseline [%; mean (SD)]	39.5 (43.9)	49.3 (38.7)	53.3 (36.3)	<0.001 (0.020)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.912	-1.419	-1.722	
PSEQ≤30 at …baseline [n (%)]	917 (37.8)	1384 (53.7)	2743 (65.1)	<0.001 (0.224)
…end of week three [n (%)]	174 (7.2)	204 (7.9)	405 (9.6)	0.001 (0.038)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.789	-1.141	-1.4	

Changes in fear-avoidance beliefs (FABQ-W, FABQ-PA), functional capacity (HFAQ), and pain self-efficacy (PSEQ) from baseline to week three. Data are presented as mean (SD) or n (%). Between-group differences were evaluated using ANCOVA with center as covariate. Partial eta-squared is reported for continuous between-group comparisons (ANCOVA), Cramer's V for categorical between-group comparisons, and Cohen's d for within-group pre-post changes.

Table 11: Changes in pharmacological treatment patterns from baseline to week three.

	CPGS I/II	CPGS III	CPGS IV	Between cohort significance (effect size: partial eta-squared / Cramer V)
Number [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	-
Simultaneous background drug treatments at … baseline [n; mean (SD)]	4.5 (1.5)	4.4 (1.5)	4.5 (1.6)	0.129 (0.000)
…end of week three [n; mean (SD)]	1.9 (1.2)	1.8 (1.2)	1.8 (1.2)	<0.001 (0.002)
Within cohort significance	<0.001	<0.001	<0.001	

(effect size; Cohen´s d)	-1.914	-2.167	-1.909	
Any pharmacological pain treatment at …baseline [n (%)]	2424 (100.0)	2579 (100.0)	4214 (100.0)	1.000 (0.000)
…end of week three [n (%)]	2144 (88.4)	2242 (86.9)	3707 (88.0)	0.237 (0.018)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.511	-0.548	-0.523	
Nonopioid analgesics at …baseline [n (%)]	1022 (42.2)	1027 (39.8)	1741 (41.3)	0.229 (0.018)
…end of week three [n (%)]	374 (15.4)	361 (14.0)	586 (13.9)	0.199 (0.019)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.618	-0.609	-0.644	
Non-steroidal antiinflammatory drugs at …baseline [NSAIDs; n (%)]	1752 (72.3)	1901 (73.7)	3168 (75.2)	0.032 (0.027)
…end of week three [n (%)]	916 (37.8)	963 (37.3)	1669 (39.6)	0.125 (0.021)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.739	-0.786	-0.771	
Mild opioid analgesics at …baseline [MOA; n (%)]	1136 (46.9)	1179 (45.7)	1859 (44.1)	0.084 (0.023)
…end of week three [n (%)]	501 (20.7)	490 (19.0)	746 (17.7)	0.012 (0.031)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.576	-0.596	-0.596	
Strong opioid analgesics at …baseline [SOA; n (%)]	628 (25.9)	640 (24.8)	1124 (26.7)	0.238 (0.018)
…end of week three [n (%)]	280 (11.6)	275 (10.7)	483 (11.5)	0.523 (0.012)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.374	-0.377	-0.395	
Cannabinoids at …baseline [CAN; n (%)]	221 (9.1)	241 (9.3)	394 (9.3)	0.945 (0.003)
…end of week three [n (%)]	109 (4.5)	128 (5.0)	219 (5.2)	0.448 (0.013)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.184	-0.171	-0.16	
Muscle relaxants at …baseline [MREL; n (%)]	1786 (73.7)	1867 (72.4)	3162 (75.0)	0.052 (0.025)
…end of week three [n (%)]	1062 (43.8)	1076 (41,7)	1778 (42.2)	0.285 (0.016)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.707	-0.652	-0.707	
Antidepressant drugs at …baseline [ADD; n (%)]	1765 (72.8)	1791 (69.4)	2970 (70.5)	0.027 (0.028)
…end of week three [n (%)]	871 (35.9)	826 (32.0)	1425 (33.8)	0.014 (0.030)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.797	-0.807	-0.789	
Antiepileptic drugs at …baseline [AED; n (%)]	607 (25.0)	570 (22.1)	921 (21.9)	0.008 (0.033)
…end of week three [n (%)]	339 (14.0)	310 (12.0)	475 (11.3)	0.005 (0.034)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-0.282	-0.27	-0.288	
Any rescue drug treatment at …baseline [n (%)]	1170 (48.3)	1288 (49.9)	2020 (47.9)	0.258 (0.017)
…end of week three [n (%)]	184 (7.6)	132 (5.1)	208 (4.9)	<0.001 (0.049)
Within cohort significance	<0.001	<0.001	<0.001	
(effect size; Cohen´s d)	-1.017	-1.16	-1.117	

Changes in background and rescue pharmacological treatments from baseline to week three across CPGS severity groups. Continuous variables (number of concurrent medications) are presented as mean (SD); categorical treatment exposure as n (%). Between-group comparisons were performed using ANCOVA (continuous outcomes) and adjusted chi-square tests (categorical outcomes), controlling for center. Partial eta-squared is reported for continuous between-group comparisons (ANCOVA), Cramer's V for categorical between-group comparisons, and Cohen's d for within-group pre-post changes.

Changes in Pain Intensity

Treatment-associated changes in pain intensity across the three predefined severity strata are summarized in Table 7. Across all severity groups, statistically significant reductions were observed for all three pain intensity measures assessed using the 100-mm visual analogue scale (VAS): lowest pain intensity within the previous 24 hours (LPI), average pain intensity (API), and highest pain intensity (HPI). Baseline pain intensity increased with CPGS severity across all three measures. After three weeks of treatment, mean pain intensity values decreased markedly in all groups. For lowest pain intensity (LPI), mean values declined from 17.9 mm in CPGS I/II, 23.1 mm in CPGS III, and 26.2 mm in CPGS IV to 9.6 mm, 13.0 mm, and 15.7 mm, respectively. Corresponding absolute reductions were 8.3 mm, 10.1 mm, and 10.5 mm (all $p < 0.001$), reflecting moderate effect sizes. For average pain intensity (API), mean baseline values of 42.5 mm, 51.8 mm, and 55.2 mm decreased to 27.5 mm, 34.3mm, and 37.3 mm in CPGS I/II, III, and IV patients, respectively. Mean reductions ranged from 15.0mm to 17.9 mm across groups (all $p < 0.001$) and were associated with large within-group effect sizes. Highest pain intensity (HPI) showed a comparable pattern. Baseline values of 66.2 mm, 75.0 mm, and 79.0 mm decreased to 48.7 mm, 57.1 mm, and 61.6 mm after treatment, corresponding to mean reductions of approximately 17–18 mm across severity strata (all $p < 0.001$) and large within-group effect sizes. Between-group comparisons of change scores revealed statistically significant differences across CPGS strata for all three pain intensity measures (ANCOVA $p < 0.001$). However, the magnitude of these between-group effects was small compared with the substantial treatment-associated improvements observed within each severity group. Responder analyses based on predefined thresholds further confirm clinically meaningful pain reductions across the cohort. Using criteria of ≥20 mm absolute reduction or ≥50% relative improvement, substantial proportions of patients achieved clinically relevant improvements in pain intensity across all severity strata.

Changes in General Well-Being and Health-Related Quality of Life

Treatment-associated changes in general well-being and health-related quality of life are summarized in Table 8. General well-being assessed with the Marburg Questionnaire on Habitual Health Findings (MQHHF; range 0–35, higher scores indicating better well-being) improved significantly across all CPGS severity strata. Mean MQHHF scores increased from 15.2 to 19.3 in CPGS I/II patients, from 1.6 to 16.4 in CPGS III patients, and from 9.5 to 14.2 in CPGS IV patients, corresponding to mean improvements of approximately 4–5 points across groups. All within-group changes were statistically significant ($p < 0.001$) and associated with moderate to large effect sizes. Physical health-related quality of life measured by the norm-based VR-12 physical component score (PCS; population mean =50) also improved significantly during the treatment period. Mean PCS values increased from 35.3 to 39.3 in CPGS I/II patients, from 30.7 to 34.9 in CPGS III patients, and from 27.4 to 31.2 in CPGS IV patients (all $p < 0.001$). Mental health-related quality of life assessed with the VR-12 mental component score (MCS) showed a comparable pattern. Mean MCS values increased from 43.2 to 47.0 in CPGS I/II patients, from 38.7 to 42.9 in CPGS III patients, and from 38.0 to 41.7 in CPGS IV patients (all $p < 0.001$). Between-group comparisons of change scores demonstrated statistically significant differences across CPGS severity strata for MQHHF, VR12-PCS, and VR12-MCS (ANCOVA $p < 0.001$). Despite statistically significant differences between severity strata, the overall pattern of results was characterized by consistent improvements across all groups, with only minor variation in the magnitude of change between strata.

Changes in Psychological Distress

Treatment-associated changes in psychological distress, assessed using the Depression Anxiety Stress Scales (DASS-21), are summarized in Table 9. Across all CPGS severity strata, statistically significant reductions were observed for depressive symptoms (DASS-21-D), anxiety (DASS-21-A), and stress (DASS-21-S) between baseline and week 3. In CPGS I/II patients, mean depression scores decreased from 5.1 to 3.4, anxiety scores from 3.7 to 2.6, and stress scores from 7.3 to 5.5 (all $p < 0.001$). Among patients classified as CPGS III, mean depression scores declined from 6.7 to 4.4, anxiety scores from 4.3 to 3.0, and stress scores from 9.1 to 6.9 (all $p < 0.001$). In the CPGS IV group, mean depression scores decreased from 7.4 to 5.1, anxiety scores from 4.6 to 3.3, and stress scores from 9.1 to 7.0 (all $p < 0.001$). Between-group comparisons of change scores showed statistically significant differences across CPGS strata for all three DASS-21 subscales (ANCOVA $p < 0.001$). While statistical comparisons between severity strata indicated modest differences in change scores, these differences were small relative to the overall magnitude of treatment-associated improvements observed across the cohort.

Changes in Functional Capacity, Fear-Avoidance Beliefs, and Pain Self-Efficacy

Treatment-associated changes in functional capacity, fear-avoidance beliefs, and pain-related self-efficacy are summarized in Table 10. Functional capacity assessed using the Hannover Functional Questionnaire Backache (HFQB; range 0–100, higher scores indicating better functional capacity) improved significantly across all CPGS severity strata. Mean HFQB scores increased from 68.5 to 77.3 in CPGS I/II patients, from 58.4 to 69.2 in CPGS III patients, and from 50.7 to 61.6 in CPGS IV patients (all $p < 0.001$), corresponding to improvements of approximately 9–11 points across groups. Fear-avoidance beliefs also decreased during the treatment period. For work-related fear avoidance (FABQ-W; range 0–42), mean scores declined from 13.8 to 10.4 in CPGS I/II patients, from 18.3 to 13.9 in CPGS III patients, and from 20.4 to 16.1 in CPGS IV patients (all $p < 0.001$). A similar pattern was observed for physical activity–related fear avoidance (FABQ-PA; range 0–24), with scores decreasing from 13.1 to 10.9, from 15.2 to 12.6, and from 15.8 to 13.4 across the respective severity strata (all $p< 0.001$). Pain self-efficacy assessed with the Pain Self-Efficacy Questionnaire (PSEQ; range 0–60) increased significantly in all groups. Mean PSEQ scores rose from 34.7 to 41.2 in CPGS I/II patients, from 29.9 to 36.7 in CPGS III patients, and from 26.6 to 33.4 in CPGS IV patients (all $p < 0.001$), corresponding to improvements of approximately 6–7 points. Improvements in pain-related self- efficacy were observed across all baseline severity strata, including patients classified as

CPGS IV, indicating that treatment-associated changes were not limited to symptom reduction but also involved patient-related coping resources relevant for the management of persistent pain. Between-group comparisons of change scores showed statistically significant differences across CPGS severity strata for functional capacity, fear-avoidance beliefs, and pain self-efficacy (ANCOVA $p < 0.001$). However, the magnitude of these between-group differences remained small relative to the consistent within-group improvements observed across all severity strata.

Changes in Pharmacological Treatment

Changes in pharmacological pain treatment between baseline (T0) and week 3 (T3) are summarized in Table 11. Across all CPGS severity strata, modest reductions were observed in the use of several medication classes during the treatment period. The proportion of patients receiving non-opioid analgesics and non-steroidal anti-inflammatory drugs (NSAIDs) decreased slightly between baseline and week 3 across all severity groups. A similar pattern was observed for opioid analgesics. Both weak and strong opioid use declined modestly during the observation period across all CPGS strata. Reductions were also observed for several adjuvant medication classes. The proportion of patients treated with muscle relaxants decreased slightly during the treatment period, with comparable modest declines observed for antidepressants and antiepileptic drugs used as co-analgesic medications. In contrast, the proportion of patients receiving cannabinoid-based medications remained largely stable between baseline and week 3. Overall, the mean number of pain-related medications per patient showed a slight decrease across all CPGS severity strata during the treatment period.

Discussion

Principal Findings

In this large real-world cohort of patients with persistent spinal pain referred for interdisciplinary second-opinion assessment prior to planned spinal surgery, intensive outpatient multimodal pain treatment was associated with substantial and clinically meaningful short-term improvements across multiple biopsychosocial outcome domains.

Among more than 13,000 patients evaluated within the structured second-opinion process, the indication for spinal surgery was not confirmed in most cases. A total of 9,217 patients subsequently entered the interdisciplinary outpatient multimodal treatment program and were included in the present analysis, representing one of the largest real-world cohorts reported to date for this type of intervention.

Pain-related disability, defined as the primary outcome of this study, decreased markedly across all baseline severity strata. Importantly, these improvements were not limited to patients with moderate symptom burden. Even patients classified as CPGS IV, representing the most severely affected subgroup with the highest levels of pain intensity and disability, demonstrated large reductions in disability and high responder rates. In several analyses, the magnitude of improvement was greatest in this subgroup. Notably, these improvements were achieved within an outpatient treatment setting and were observed consistently across a large nationwide network of interdisciplinary pain centers.

Beyond the primary endpoint, consistent treatment-associated improvements were observed across a broad range of secondary outcomes reflecting the multidimensional burden of persistent spinal pain. Significant improvements were documented for pain intensity, functional capacity, psychological distress, health-related quality of life, general well-being, and pain-related self-efficacy. Multivariate analyses confirmed the robustness of these findings, demonstrating consistent patterns of change across outcome domains while indicating no relevant center-related effects across participating interdisciplinary treatment centers. Moreover, multivariate findings remained robust after adjustment for demographic covariates, indicating that baseline severity differences in treatment- associated change patterns were not explained by age or sex differences between groups.

Taken together, these results indicate that an intensive interdisciplinary outpatient multimodal treatment program can achieve substantial short-term improvements across multiple biopsychosocial imensions of persistent spinal pain, even in patients with high baseline disability and complex clinical presentations. Importantly, these findings should be interpreted within the context of the structured interdisciplinary second-opinion pathway in which the present treatment program is embedded. For many patients with persistent spinal pain, this pathway may provide an evidence-informed conservative treatment alternative at a critical decision point where surgical intervention might otherwise be considered.

Multivariate analyses further demonstrated that baseline severity remained a significant determinant of treatment-associated change patterns across outcome domains even after adjustment for demographic covariates and treatment center. The observed multivariate effect size was in the moderate range (partial $\eta^2 \approx 0.06$), which is notable in the context of a large real-world multicenter cohort characterized by substantial clinical heterogeneity. At the same time, the absence of relevant center effects suggests that treatment-associated improvements were highly consistent across participating centers within the network.

Position Within the Current Literature

Interdisciplinary multimodal pain treatment is widely regarded as a key therapeutic approach for patients with persistent spinal pain who do not respond adequately to unimodal interventions. Current clinical guidelines as well as a substantial body of empirical research support the use of multimodal treatment strategies that integrate medical, physical, and psychological interventions to address the complex biopsychosocial mechanisms underlying chronic pain. Several systematic reviews and meta-analyses have demonstrated that multidisciplinary biopsychosocial rehabilitation programs can lead to clinically meaningful improvements in pain intensity, functional capacity, and health- related quality of life in patients with chronic low back pain [10,11,61].

However, much of the available evidence on multimodal pain therapy has been generated in inpatient treatment settings. In several European health care systems, including Germany,

specialized inpatient multimodal pain therapy programs have historically represented the dominant model for delivering intensive interdisciplinary pain treatment [62]. Consequently, comparatively little evidence is available regarding the effectiveness of structured intensive outpatient multimodal treatment programs.

This evidence gap is particularly relevant in the context of ongoing health system efforts to strengthen outpatient care structures and to reduce reliance on inpatient treatment whenever clinically feasible [63,64]. Structured outpatient programs may offer the opportunity to provide intensive interdisciplinary treatment while maintaining patients within their everyday social and occupational environments, which may facilitate functional rehabilitation and the application of coping strategies in real-life contexts [14]. Although several studies have investigated outpatient multidisciplinary pain management programs, most published cohorts have been considerably smaller and often limited to single-center settings [11,12,61].

The present study contributes to the existing literature by providing outcome data from a large real-world cohort treated within a nationwide interdisciplinary outpatient treatment network. The observed improvements across multiple clinical domains are consistent with the established evidence base suppor τ ng multimodal pain therapy and suggest that comparable clinical benefits can also be achieved in well-structured outpatient treatment programs. These findings therefore extend the current evidence base by demonstrating the feasibility and effectiveness of intensive multimodal pain treatment in a large-scale outpatient care setting.

Role of the Interdisciplinary Second-Opinion Pathway

The findings of the present study should also be interpreted in the context of the interdisciplinary second-opinion pathway in which the treatment program is embedded. Previous studies have demonstrated substantial variability in the indication for spinal surgery, with marked regional differences in surgical rates and considerable uncertainty regarding the expected benefit for individual patients [17,65]. In response to these challenges, second-opinion procedures have been introduced in several health care systems as a strategy to improve the appropriateness of surgical indications and to promote guideline-concordant conservative management when appropriate [66,67].

Within such frameworks, interdisciplinary assessment may play a particularly important role by broadening the evaluation beyond structural imaging findings and by incorporating functional, psychological, and social dimensions of persistent spinal pain. The results of the present study suggest that structured second-opinion pathways may not only help prevent potentially unnecessary surgery but can also facilitate timely access to interdisciplinary conservative treatment options for patients with complex and disabling pain conditions.

Clinical Interpretation

The present findings challenge the widely held assumption that patients with severe persistent spinal pain and pronounced disability require treatment within inpatient multimodal pain programs in order to achieve meaningful clinical improvement.

In many health care systems, including Germany, patients with high levels of pain-related disability are frequently considered unsuitable for outpatient multimodal treatment because of the perceived complexity and severity of their clinical condition. Consequently, intensive multidisciplinary pain treatment has traditionally been associated primarily with inpatient care settings.

The results of the present real-world evaluation suggest that the effectiveness of multimodal pain therapy may depend less on the treatment setting itself and more on the structural characteristics of the treatment program. In the network evaluated in this study, patients received coordinated medical, physiotherapeutic, and psychological interventions delivered within a structured treatment schedule while maintaining individualized treatment planning based on the specific needs of each patient.

Under these conditions, substantial and clinically meaningful improvements were observed even in patients with very high baseline disability and long-standing pain conditions. Importantly, treatment responses were not confined to moderately affected patients but were also pronounced among individuals classified as CPGS IV, representing the most severely affected subgroup in this cohort. These observations indicate that well-structured interdisciplinary outpatient multimodal treatment programs can provide effective care for a broad spectrum of patients with persistent spinal pain, including individuals who are often considered candidates for inpatient treatment.

Severity of the Study Population and Clinical Relevance

An important aspect when interpreting the present findings is the high clinical complexity of the study population. Patients included in this analysis underwent an interdisciplinary second-opinion assessment prior to a planned spinal surgical intervention. Access to the second-opinion program occurred through multiple pathways, including direct patient self-referral, referral by treating physicians, or mediation by statutory health insurance providers. Thus, the cohort represents individuals for whom surgical treatment had already been considered within routine care. Consistent with this clinical context, patients in the present cohort showed substantial baseline disease burden cross multiple domains. High levels of pain intensity, pronounced pain-related disability, marked functional limitations, and considerable psychological distress were common. In addition, most patients had undergone extensive previous treatment attempts, including pharmacological therapies, interventional procedures, and, in a substantial proportion of cases, prior spine surgery.

These characteristics reflect a patient population that is often regarded as particularly difficult to treat within conventional outpatient care structures. In many clinical settings, such patients are therefore considered potential candidates for escalation of treatment intensity, including surgical intervention or referral to specialized inpatient multimodal pain programs.

Against this background, the magnitude and consistency of the improvements observed in the present study are of clinical relevance. Even within this highly burdened population, clinically

meaningful improvements were observed across multiple domains of pain-related health status following completion of the structured interdisciplinary outpatient treatment program.

Role of Pain Self-Efficacy

A further relevant finding concerns the observed improvements in pain-related self-efficacy. Effect sizes for the Pain Self-Efficacy Questionnaire were large across all severity strata and were highest among patients classified as CPGS IV. Self-efficacy beliefs are considered a key psychological construct within the biopsychosocial model of chronic pain and are closely linked to functional outcomes and long-term adaptation to persistent pain conditions [49,50]. In the present study, substantial improvements in pain self-efficacy were observed across all baseline severity strata, including patients with the highest levels of disability. This finding is clinically relevant because increased self-efficacy has been associated with improved coping strategies, greater engagement in active self-management, and more favorable long-term functional trajectories in patients with chronic pain. Improvements in this domain may therefore represent an important mechanism through which interdisciplinary multimodal treatment contributes to sustained clinical benefit beyond the immediate treatment phase.

Implications for Health Care Delivery

The findings of the present study have potential implications for the organization of care for patients with persistent spinal pain. Although interdisciplinary multimodal pain therapy is widely recognized as an effective treatment approach for complex chronic pain conditions, access to such programs remains limited in many health care systems. In particular, inpatient multimodal pain programs are resource-intensive and typically concentrated in specialized centers with restricted treatment capacity. At the same time, structural developments in several health care systems may further constrain the availability of inpatient treatment options. In this context, identifying effective outpatient treatment models becomes increasingly important to ensure continued access to interdisciplinary pain care.

The present analysis suggests that intensive interdisciplinary outpatient multimodal treatment programs may represent a viable and scalable alternative for many patients with persistent spinal pain. Within the structured network evaluated in this study, clinically meaningful improvements were observed across multiple outcome domains despite high baseline levels of disability and extensive prior treatment exposure. These findings indicate that interdisciplinary multimodal pain therapy does not necessarily require an inpatient setting to achieve meaningful clinical benefit.

An additional structural feature of the care model evaluated in this study is the use of a bundled reimbursement approach rather than a traditional procedure-based fee-for-service system. This payment structure allows flexible allocation of therapeutic interventions across disciplines and may facilitate coordinated interdisciplinary care delivery in outpatient settings.

From a health system perspective, structured outpatient multimodal programs may therefore represent an important component of stepped-care approaches to pain management. By delivering coordinated multidisciplinary treatment within an outpatient framework, such programs may expand access to evidence-based care while potentially reducing reliance on hospital-based treatment pathways.

Importantly, outpatient multimodal treatment is not intended to replace inpatient programs entirely. Rather, it may function as an intermediate level within a continuum of care, allowing many patients with persistent spinal pain to receive intensive interdisciplinary treatment without requiring hospital admission while preserving inpatient resources for the most complex cases. The nationwide treatment network evaluated in this study further suggests that such outpatient models can be implemented across diverse clinical settings while maintaining consistent treatment outcomes. These findings may be particularly relevant in health systems facing increasing demand for chronic pain treatment and limited inpatient capacity.

Patients Who May Require Inpatient Treatment

Although the results of the present analysis demonstrate substantial improvements across all severity strata including patients with very high levels of baseline disability—it is unlikely that outpatient multimodal treatment represents the optimal treatment setting for every patient with persistent spinal pain.

Multimodal pain therapy addresses complex biopsychosocial interactions between physical symptoms, psychological factors, and functional limitations. In some patients, the severity of these interactions may require treatment environments that allow continuous multidisciplinary supervision and more intensive therapeutic structures than can typically be provided in outpatient settings. Patients with extremely severe functional impairment, pronounced psychiatric comorbidity, or unstable social circumstances may therefore still benefit from treatment within specialized inpatient pain programs. Similarly, patients requiring complex medication adjustments, intensive psychological stabilization, or highly coordinated therapeutic interventions may require treatment settings that provide continuous interdisciplinary supervision.

The findings of the present study should therefore not be interpreted as suggesting that inpatient multimodal pain therapy is no longer necessary. Rather, they indicate that many patients with persistent spinal pain—including individuals with high baseline disability—can achieve meaningful improvements within well-structured interdisciplinary outpatient treatment programs. Within stepped-care models of pain management, outpatient multimodal treatment may therefore represent an important intermediate treatment level, reserving inpatient programs for patients with the highest levels of clinical complexity or for individuals who do not respond sufficiently to intensive outpatient approaches.

The present study did not specifically investigate predictors of insufficient treatment response. Identifying patient characteristics associated with limited benefit from outpatient multimodal programs represents an important objective for future research and may help refine clinical decision-making regarding optimal

allocation to outpatient versus inpatient treatment pathways.

Strengths and Limitations

The present study has several important strengths that should be considered when interpreting the findings. The analysis is based on a very large real-world cohort comprising more than nine thousand patients treated within an interdisciplinary outpatient multimodal pain treatment network. To our knowledge, this represents one of the largest observational evaluations of interdisciplinary outpatient multimodal pain therapy reported to date. The large sample size allowed detailed stratified analyses across different baseline severity groups, including patients with the highest levels of pain-related disability and enabled the use of multivariate statistical approaches to examine treatment-associated change patterns across multiple outcome domains.

Another important strength is that the study reflects routine clinical practice rather than highly selected experimental treatment settings. All patients were treated within a structured network of specialized pain treatment centers operating under real-world conditions. The findings therefore provide insight into treatment outcomes achievable in everyday interdisciplinary pain care rather than under highly controlled experimental circumstances. In this sense, the present investigation can be considered a pragmatic effectiveness evaluation of interdisciplinary outpatient multimodal pain treatment within routine clinical care.

In addition, the study employed a comprehensive multidimensional outcome assessment reflecting the biopsychosocial model of chronic pain. Outcomes included measures of pain intensity, pain-related disability, functional capacity, psychological distress, health-related quality of life, and pain-related self-efficacy. The use of validated patient-reported outcome measures across multiple domains allowed a broad evaluation of treatment-associated changes beyond pain intensity alone.

The clinical relevance of the cohort represents another important strength of the present analysis. Treatment outcomes were evaluated in a population of patients with substantial disease burden and extensive prior treatment exposure. Many patients had already received multiple pharmacological, interventional, and rehabilitative treatment approaches before referral to the interdisciplinary second-opinion assessment. Demonstrating clinically meaningful improvements in such a heavily pretreated population strengthens the relevance of the observed treatment effects.

Finally, the high completeness of the dataset reflects the structured documentation procedures implemented within the German Pain e-Registry, including predefined assessment schedules and automated plausibility checks during data entry, thereby minimizing missing data and enhancing the robustness of the present analysis.

Several limitations should also be considered. First, the analysis was observational and did not include a randomized control group. Consequently, causal inferences regarding treatment effects should be made with caution. However, the magnitude and consistency of the observed improvements across multiple outcome domains, the large cohort size, and the comparable treatment-associated changes observed across predefined severity strata suggest that the findings are unlikely to be explained solely by natural fluctuations in symptom severity.

Second, outcome assessment focused on short-term treatment effects during the three-week intervention period. Although this timeframe corresponds to the active treatment phase of the program, the long-term sustainability of the observed improvements cannot be evaluated within the present analysis. Follow-up data collected at later time points will be required to determine whether treatment-associated benefits persist over time.

Third, the study population consisted of patients referred to an interdisciplinary second opinion prior to a planned spinal surgical intervention. The cohort therefore represents a specific subgroup of patients with persistent spinal pain characterized by high symptom burden and extensive prior treatment exposure. While this increases the clinical relevance of the findings, it may limit the generalizability of the results to patients with less complex pain conditions.

Fourth, although treatment was delivered within a structured interdisciplinary program, specific therapeutic components were individualized according to the needs of each patient. While this reflects the principles of patient-centered multimodal pain therapy in routine clinical practice, it limits the ability to determine which particular treatment components contributed most strongly to the observed improvements.

Finally, although treatment was delivered across a large number of participating centers, the analyses did not indicate relevant center-related influences on treatment outcomes. Nevertheless, as with all multicenter real-world evaluations, unmeasured differences in local treatment implementation cannot be entirely excluded.

These strengths and limitations should be considered when interpreting the findings of this real-world evaluation of interdisciplinary outpatient multimodal pain treatment.

Conclusions

In this large real-world cohort of patients with persistent spinal pain referred for interdisciplinary second-opinion assessment prior to planned spinal surgery, intensive interdisciplinary outpatient multimodal treatment was associated with substantial short-term improvements across multiple clinically relevant outcome domains.

Clinically meaningful benefits were observed not only in patients with moderate symptom burden but also in those with the highest levels of baseline disability. Improvements were also observed in pain-related self-efficacy, a key factor associated with active coping and long-term adaptation to persistent pain conditions, suggesting that effective interdisciplinary pain treatment can be achieved in appropriately structured outpatient settings.

Within stepped-care models of pain management, intensive outpatient multimodal programs may therefore represent an important treatment level, enabling many patients with complex

persistent spinal pain to receive coordinated interdisciplinary care without requiring hospital admission.

Further research should examine the long-term sustainability of treatment effects and help define patient characteristics that guide optimal allocation between outpatient and inpatient multimodal pain therapy.

Acknowledgement

We would like to express our sincere gratitude to all members of our interdisciplinary IMC network for their exceptional dedication, professionalism, and ongoing commitment to the implementation of this demanding yet highly effective IMC outpatient interdisciplinary multimodal pain therapy program. In particular, we would like to thank the physicians, physiotherapists, psychotherapists, occupational therapists, medical assistants, pain nurse specialists/ algesiological assistants, and all other colleagues involved in patient care and program coordination. Their collaborative spirit, expertise, and daily engagement were essential to the success of this program and to the outcomes presented in this work.

Conflict of Interest

The authors are founders and owners of IMC GmbH and developers of the interdisciplinary multimodal pain management program presented in this article. Nevertheless, the authors declare that they have no competing interests or conflicts of interest regarding the reporting and publication of the results presented in this study.

References

1. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, et al. (2018) What low back pain is and why we need to pay attention. Lancet 391(10137): 2356-2367.
2. Buchbinder R, van Tulder M, Öberg B, Costa LM, Woolf A, et al. (2018) Low back pain: a call for action. Lancet 391(10137): 2384-2388.
3. GBD 2019 Diseases and Injuries Collaborators (2020) Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396(10258):1204-1222. Erratum in: Lancet 396(10262): 1562.
4. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW (1990) Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am 72(3): 403-8.
5. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, et al. (1994) Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 331(2): 69-73.
6. Brinjikji W, Luetmer PH, Comstock B, Bresnahan BW, Chen LE, et al. (2015) Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol 36(4): 811-6.
7. (2020) Low back pain and sciatica in over 16s: assessment and management. London: National Institute for Health and Care Excellence (NICE) (NICE Guideline, No. 59.).
8. Qaseem A, Wilt TJ, McLean RM, Forciea MA; Clinical Guidelines Committee of the American College of Physicians, et al. (2017) Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. Ann Intern Med 166(7): 514-530.
9. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) (2017) Nationale VersorgungsLeitlinie Nicht-spezifischer Kreuzschmerz – Langfassung, 2. Auflage. Version 1. [accessed online: March 5, 2026].
10. Guzmán J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, et al. (2001) Multidisciplinary rehabilitation for chronic low back pain: systematic review. BMJ 322(7301): 1511- 6.
11. Scascighini L, Toma V, Dober-Spielmann S, Sprott H (2008) Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes. Rheumatology (Oxford) 47(5): 670-8.
12. Gatchel RJ, McGeary DD, McGeary CA, Lippe B (2014) Interdisciplinary chronic pain management: past, present, and future. Am Psychol 69(2): 119-30.
13. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RW, et al. (2014) Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2014(9): CD000963.
14. Waddell G (1987) 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low- back pain. Spine (Phila Pa 1976) 12(7): 632-44.
15. Turk DC, Okifuji A (2002) Psychological factors in chronic pain: evolution and revolution. J Consult Clin Psychol 70(3): 678-90.
16. Lurie JD, Weinstein JN (2001) Shared decision-making and the orthopaedic workforce. Clin Orthop Relat Res (385): 68-75.
17. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES (2006) United States' trends and regional variations in lumbar spine surgery: 1992-2003. Spine (Phila Pa 1976) 31(23): 2707.
18. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, et al. (2012) Shared decision making: a model for clinical practice. J Gen Intern Med 27(10): 1361-7.
19. Von Korff M, Ormel J, Keefe FJ, Dworkin SF (1992) Grading the severity of chronic pain. Pain 50(2): 133-149.
20. Smith BH, Penny KI, Purves AM, Munro C, Wilson B, et al. (1997) The Chronic Pain Grade questionnaire: validation and reliability in postal research. Pain 71(2): 141-7.
21. Dunn KM, Croft PR (2005) Classification of low back pain in primary care: using "bothersomeness" to identify the most severe cases. Spine (Phila Pa 1976) 30(16):1887-92.
22. Dunn KM, Jordan K, Croft PR (2006) Characterizing the course of low back pain: a latent class analysis. Am J Epidemiol 163(8):754-61.
23. Pollard CA (1984) Preliminary validity study of the pain disability index. Percept Mot Skills 59(3):974.
24. Tait RC, Chibnall JT, Krause S (1990) The Pain Disability Index: psychometric properties. Pain 40(2):171-182.
25. Soer R, Köke AJ, Vroomen PC, Stegeman P, Smeets RJ, Coppes MH, Reneman MF (2013) Extensive validation of the pain disability index in 3 groups of patients with musculoskeletal pain. Spine (Phila Pa 1976) 38(9): E562-8.
26. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole MR (2001) Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 94(2): 149-158.
27. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, et al. (2008) Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 9(2):105-21.
28. Cleeland CS, Ryan KM (1994) Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singap 23(2): 129-38.
29. Huskisson EC (1974) Measurement of pain. Lancet 2(7889): 1127-31.
30. Hawker GA, Mian S, Kendzerska T, French M (2011) Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short

Form- 36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 63 Suppl 11: S240-52.

31. Kelly AM (2001) The minimum clinically significant difference in visual analogue scale pain score doesnot differ with severity of pain. Emerg Med J 18(3): 205-7.

32. Kohlmann T, Raspe H (1996) Der Funktionsfragebogen Hannover zur alltagsnahen Diagnostik der Funktionsbeeinträchtigung durch Rückenschmerzen (FFbH-R) [Hannover Functional Questionnaire in ambulatory diagnosis of functional disability caused by backache]. Rehabilitation (Stuttg) 35(1): I-VIII. German.

33. Roese I, Kohlmann T, Raspe H (1996) Zur Messung der Funktionskapazität bei Rückenschmerzpatienten in der Rehabilitation: ein Vergleich standardisierter Fragebogen [Measuring functional capacity in backache patients in rehabilitation: a comparison of standardized questionnaires]. Rehabilitation (Stuttg) 35(2):103-8. German.

34. Magnussen L, Strand LI, Lygren H (2004) Reliability and validity of the back performance scale: observing activity limitation in patients with back pain. Spine (Phila Pa 1976) 29(8): 903-7.

35. Waddell G, Newton M, Henderson I, Somerville D, Main CJ (1993) A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 52(2): 157-168.

36. Pfingsten M, Kröner-Herwig B, Leibing E, Kronshage U, Hildebrandt J (2000) Validation of the German version of the Fear-Avoidance Beliefs Questionnaire (FABQ). Eur J Pain 4(3): 259-66.

37. Vlaeyen JWS, Linton SJ (2000) Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 85(3): 317-332.

38. George SZ, Fritz JM, Erhard RE (2001) A comparison of fear-avoidance beliefs in patients with lumbar spine pain and cervical spine pain. Spine (Phila Pa 1976) 26(19):2139-45.

39. Lovibond SH, Lovibond PF (1995) The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour Research and Therapy 33(3): 335–343.

40. Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP (1998) Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychological Assessment 10(2): 176–181.

41. Henry JD, Crawford JR (2005) The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. Br J Clin Psychol 44(Pt 2): 227-39.

42. Freynhagen R, Baron R, Gockel U, Tölle TR (2006) painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Current Medical Research and Opinion 22(10): 1911–1920.

43. Hallström H, Norrbrink (2011) Screening tools for neuropathic pain: can they be of use in individuals with spinal cord injury? Pain 152(4):772-779.

44. Freynhagen R, Tölle TR, Gockel U, Baron R (2016) The painDETECT project - far more than a screening tool on neuropathic pain. Curr Med Res Opin 32(6): 1033-57.

45. Selim AJ, Rogers W, Fleishman JA, Qian SX, Fincke BG, et al. (2009) Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). Qual Life Res 18(1):43-52.

46. Ware JE, Kosinski M, Turner-Bowker DM, Gandek B (2002) How to score version 2 of the SF-12 Health Survey. Lincoln, RI: QualityMetric.

47. Basler H-D (2000) The Marburg questionnaire on habitual health findings - A study on patients with chronic pain Der Schmerz13(6): 385-91.

48. Brähler E, Schumacher J, Strauß B (2002) Diagnostische Verfahren in der Psychotherapie. Göttingen: Hogrefe.

49. Nicholas MK (2007) The pain self-efficacy questionnaire: Taking pain into account. Eur J Pain (2): 153-63.

50. Nicholas MK, Asghari A, Blyth FM (2008) What do the numbers mean? Normative data in chronic pain measures. Pain 134(1-2): 158-73.

51. Bandura A (1977) Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 84(2): 191-215.

52. Ruehlman LS, Karoly P, Newton C, Aiken LS (2005) The development and preliminary validation of a brief measure of chronic pain impact for use in the general population. Pain 113(1-2): 82-90.

53. Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 4: 863.

54. Cohen J (1988) Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

55. Altman DG (1991) Practical Statistics for Medical Research. London: Chapman & Hall: 56.

56. Vickers AJ, Altman DG (2001) Statistics notes: Analysing controlled trials with baseline and follow up measurements. BMJ 323(7321): 1123-4.

57. Johnson RA, Wichern DW (2007) Applied Multivariate Statistical Analysis. 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

58. Bonferroni CE (1936) Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze.

59. Dunn OJ (1961) Multiple comparisons among means. Journal of the American Statistical Association 56(293): 52–64.

60. Box GEP (1949) A general distribution theory for a class of likelihood criteria. Biometrika 36(3–4): 317-346.

61. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RW, et al. (2015) Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. BMJ 350: h444.

62. Pfingsten M, Arnold B, Böger A, Brinkschmidt T, Casser HR, et al. (2019) Sektorenübergreifende interdisziplinäre multimodale Schmerztherapie : Empfehlungen zu Struktur- und Prozessparametern der Ad-hoc-Kommission „Interdisziplinäre Multimodale Schmerztherapie“ der Deutschen Schmerzgesellschaft e. V [Cross-sectoral interdisciplinary multimodal pain therapy : Recommendations on structural and process parameters of the ad hoc commission “Interdisciplinary Multimodal Pain Therapy” of the German Pain Society (Deutsche Schmerzgesellschaft e. V.)]. Schmerz 33(3): 191-203. German.

63. OECD (2021) Health at a Glance 2021: OECD Indicators. Paris: OECD Publishing.

64. Busse R, Blümel M, Knieps F, Bärnighausen T (2017) Statutory health insurance in Germany: a health system shaped by 135 years of solidarity, self-governance, and competition. Lancet. 2017 Aug 26;390(10097): 882-897.

65. Deyo RA, Mirza SK, Martin BI (2006) Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. Spine (Phila Pa 1976) 31(23): 2724-7.

66. Ferreira GE, Zadro J, Liu C, Harris IA, Maher CG (2022) Second opinions for spinal surgery: a scoping review. BMC Health Serv Res 22(1): 358.

67. Epstein NE (2013) Are recommended spine operations either unnecessary or too complex? Evidence from second opinions. Surg Neurol Int 4(Suppl 5): S353-8.