DEEP Study: Utility of the multidimensional pain inventory in persistent orofacial pain.

Running title: Utility of MPI in persistent orofacial pain

James R Allison[1, 2 *], Chris Penlington[1, 2], Justin Durham[1, 2]

1- School of Dental Sciences, Newcastle University, Newcastle upon Tyne, UK
2- Newcastle upon Tyne Hospitals NHS Foundation Trust

***Corresponding Author:**

Mr James Allison
School of Dental Sciences
Framlington Place
Newcastle University
Newcastle upon Tyne
NE2 4BW

James.allison@newcastle.ac.uk
0191 2087823

Acknowledgements

We thank the patients, health professionals, and steering group members for their valuable time and assistance. This study and JD are funded by the National Institute for Health Research (NIHR; Clinician Scientist Award NIHR-CS-011-003). JRA received funding for this work from a Helen Tonge Award from the School of Dental Sciences, Newcastle University.

Abstract

Background

Persistent orofacial pain (POFP) is disabling and patients' treatment outcomes are difficult to predict; psychosocial factors play a role. The West Haven-Yale Multidimensional Pain Inventory (MPI) is a self-report measure which to our knowledge, has not been studied across primary and secondary care in heterogeneous POFP.

Objective

Assess the MPI's ability to predict clinical outcome in POFP patients across primary and secondary care settings receiving usual care.

Methods

146 patients receiving usual care for POFP were recruited from primary and secondary care medical and dental practices in north east England. Participants completed the MPI (v3) and Graded Chronic Pain Scale (GCPS) at recruitment, and after 6, 12, 18, and 24 months. The Patient Health Questionnaire-4 (PHQ-4) was completed at recruitment, 12, and 24 months. "Good" and "poor" outcome status was assigned to participants based on their mode dichotomised GCPS score. Logistic regression was used with overall GCPS outcome (good/poor) as the dependent variable and MPI subscale scores, demographic variables, and PHQ-4 as predictors.

Results

110 participants had a "good", and 36 a "poor" outcome. In the "poor" outcome group: age, mean income, and life control scores were lower; deprivation, months in pain, PHQ-4, pain severity, interference, and affective distress scores were higher. In the "good" group MPI scores improved over time. Interference was the only consistent predictor of "poor" outcome in the logistic regression model (OR: 1.14 – 1.98, $p < 0.05$).

Conclusion

The MPI interference subscale may help to identify patients with POFP who are likely to have consistent pain-related disability over time; it may therefore be clinically useful to identify patients likely to need early intervention.

Keywords: Facial Pain; Chronic Pain; Longitudinal Studies; Observational Studies; Psychometrics; Prognosis

Background

Persistent orofacial pain (POFP) encompasses a number of specific diagnoses relating to painful conditions of non-dental aetiology in the face and mouth lasting longer than 3 months [1]. These conditions are estimated to affect 5-7% of the population [2], and can be very disabling for sufferers [3]. Many clinicians find the management of POFP challenging and although many patients do well with simple treatment, others obtain little improvement [4]. It can often be difficult to identify at the outset from physical diagnoses which patients will do well with conservative management, and which will not [5].

Although the pathophysiology of many types of POFP is not completely understood, for many conditions causing POFP, psychosocial factors play a significant role [6]. Various measures assessing psychosocial aspects of patients' conditions have been used to assess psychosocial factors in patients with orofacial pain. One such measure is the West Haven-Yale Multidimensional Pain Inventory (MPI) which was designed to assess pain intensity, pain related disability, psychosocial distress, social support, and levels of activity [7]. The MPI has been used in a number of patient groups, including those with general chronic pain, lower back and neck pain, fibromyalgia, and temporomandibular disorders (TMD), with good reported validity [7-9]. Version 3 of the MPI is a 48-item self-report measure comprising 9 subscales (pain severity, interference, life control, affective distress, support, punishing responses, solicitous responses, distracting responses, and general activity) [10]. The MPI has been studied in patients with orofacial pain, most commonly in patients with temporomandibular disorders (TMD) [9, 11-14], but also heterogeneous groups with acute [15] and chronic orofacial pain [12, 16], and in patients with burning mouth syndrome [17, 18]. To the authors' knowledge the measure has not been studied in a heterogeneous population of patients with POFP across both primary and secondary care settings.

The aim of the present study is to assess the utility of the MPI in patients across different settings receiving "usual care" (routine management of their condition in primary or secondary care settings according to the treating clinician) for persistent orofacial pain. Our objective is to examine whether the scale can differentiate between patients who are likely to achieve good outcomes with usual care and those who are not.

Methods

The present study was part of the larger DEEP study (Developing Effective and Efficient care pathways in chronic Pain), which was a closed cohort, longitudinal study of patients receiving routine care for POFP across primary and secondary care settings followed over 24 months. Ethical approval was obtained from the UK NHS National Research Ethics Service Committee Leeds West (Ref 12-YH-0338) and the full DEEP study protocol is reported elsewhere [19]. Briefly however, patients with any orofacial pain of greater than 3 months duration were recruited from medical and dental practices in primary care, and from neurology, oral and maxillofacial surgery, dental emergency, oral medicine, and restorative dentistry clinics in secondary care across the north east of England. Eligible patients were identified using two validated screening measures [20, 21] and gave written consent after agreeing to participate. All participants were asked to complete several study questionnaires including those described below, and these measures were repeated at 6, 12, 18 and 24 months. An *a priori* power calculation with $\alpha = 0.05$ (two-tailed), power = 0.8 and effect size $d = 0.4$ suggested a sample size of $n = 200$ would be sufficient to detect differences between two groups based on the Graded Chronic Pain Scale which was the primary outcome measure for this study. This study conforms to STROBE guidelines.

Measures and Instruments

Participants were given several study questionnaires in a standardised order over their time in the DEEP study. This included the following instruments that were of relevance to the present study:

Section 1 of the MPI [7] (version 3) which includes the Pain Severity, Interference, Life Control, Affective Distress, and Support subscales was completed at baseline, and at 6, 12, 18, and 24 months. Section 2 of the MPI, which includes the Punishing Responses, Solicitous Responses, and Distracting Responses subscales, was completed at baseline only. Section 3 of the MPI (activity subscales) was not included to reduce participant burden. Summary scores for each subscale of the MPI were calculated using the item response theory method (Rasch model of MPI version 3 software [10]).

The Graded Chronic Pain Scale (GCPS) [22] was completed by participants at recruitment, and at 6, 12, 18, and 24 months. The GCPS measures pain intensity and disability and is widely used in this population. Pain intensity and disability are rated from 0 to 10 across six questions, and the mean of these responses is combined with the number of days which the respondent has been prevented from carrying out their usual daily activities in the prior 6 months. This produces a grade from 0 to 4 which considers both the intensity of the pain, and the resulting level of disability. For each time point, the GCPS score was dichotomised into "low" (grade 0, 1, or 2a) and "high" (grade 2b, 3, or 4) as described by Dworkin *et al.* [23]. Accordingly, those with a "low" dichotomised GCPS score (dGCPS) had no disability resulting from their pain, and those with a "high" dGCPS had some level of disability.

The Patient Health Questionnaire-4 (PHQ-4) [24] is a brief screening tool which combines two items measuring depression from the more detailed PHQ-9 and two items measuring anxiety from the Generalised Anxiety Disorder-7 (GAD-7) measure. Items are scored on a four-point scale. This measure was completed by participants at recruitment, 12 months and 24 months.

'Outcome' measure (pain-related disability)

As POFP encompasses conditions which tend to vary over time, we derived an overall 'outcome' across the study period (i.e., good/poor outcome over time) from the dGCPS at all time points, which represents pain-related disability. This was obtained by calculating the mode dGCPS score from all five time points, and thus those assigned to the 'good outcome' group had a 'low' dGCPS (no pain related disability reported) on three or more occasions, and those in the 'poor outcome' group had a 'high' dGCPS on three or more occasions. As the mode dGCPS score across all five time points could not be calculated where fewer than three of the same dGCPS scores were recorded (i.e., three 'high' or three 'low' dGCPS scores), these participants were excluded. This overall GCPS outcome was used as the primary outcome measure.

Statistical methods

Data were analysed using Stata version 16 (StataCorp; Texas, USA) and SPSS version 25 (IBM Corp.; New York, USA). To assess differences between good and poor outcome groups the following statistical tests were used: chi-squared for nominal and ordinal variables; two sample *t*-tests for normally distributed continuous variables; and two-sample Wilcoxon rank-sum and Friedman tests for non-normally distributed variables. Normality was assessed graphically and using the Shapiro-Wilk test. Spearman's correlation coefficients were calculated to assess the relationship between demographic variables, GCPS score (0 – 4, non-dichotomised), PHQ-4, and each of the MPI subscales. Complete case analysis was used for these analyses.

For the primary analysis, missing data were imputed by multiple imputation using chained equations (MICE) [25]; this method was chosen as data were missing for both ordinal and continuous variables and the pattern of missing data was non-monotone. Missing values were assumed to be missing at random. Imputation was carried out using ordered logistic regression for ordinal variables (education classification, PHQ-4 scores) and linear regression for continuous variables (income, IMD score, and MPI subscale Rasch scores). The complete variables: overall GCPS outcome, gender, age, and months in pain were included in the imputation model as auxiliary variables. 10 imputations were generated, with 600 iterations in total (burn-in and burn-between = 60). The imputation model was assessed graphically using cumulative distribution functions and by comparing the means and standard deviations of imputed datasets with those of the observed data.

Logistic regression models were then fitted to the imputed datasets at each of the five time points. Overall GCPS outcome (good/poor) was the dependent variable and MPI subscale Rasch scores, demographic variables (age, gender, education, deprivation, months in pain), and PHQ-4 score and were included as predictors.

Results

201 participants were recruited to the DEEP study and dropout at 24 months was 34.33% as reported in detail elsewhere [26]. 198 participants returned GCPS and MPI questionnaires at baseline, 172 were returned at 6 months, 156 at 12 months, 137 at 18 months, and 131 at 24 months. 146 participants had at least three high or three low dGCPS scores over the five time points, allowing calculation of an overall GCPS outcome (good/poor). These 146 participants were included in the present analysis. Missing data were below 10% for all variables except monthly income, PHQ-4 score at 24 months, and MPI score at 24 months (25.3%, 19.9% and 11% respectively); missingness for each variable is shown in table 1. There were no missing data for the outcome variable, overall GCPS outcome.

Demographic characteristics

110 participants had a good overall GCPS outcome ('low' mode dGCPS score across all 5 time points), and 36 had a poor overall GCPS outcome ('high' mode dGCPS score across all 5 time points). Overall distribution of GCPS scores are presented in the Supplementary Data (online). Mean age across all participants was 53.3 years (SD = 14.6), and this was significantly lower in the poor outcome group ($t(144) = 2.99$, $p < 0.01$). 82.2% of all participants were female, this proportion did not differ between good and poor outcome groups ($\chi^2(1, N=146) = 0.09$, $p = 0.77$). Mean Index of Multiple Deprivation (IMD) scores were significantly higher in the poor outcome group ($Z = -3.00$, $p < 0.01$), indicating a greater degree of deprivation. Mean number of months in pain at baseline was significantly higher in the poor outcome group ($Z = -2.01$, $p = 0.04$), as was PHQ-4 score at baseline ($Z = -3.56$, $p < 0.01$), 12 months ($Z = -4.64$, $p < 0.01$), and 24 months ($Z = -5.20$, $p < 0.01$). Mean income was higher in the good outcome group ($Z = 2.11$, $p = 0.03$) and level of education did not differ between groups (χ^2 (5, N=132) = 7.49, $p = 0.19$). Demographic characteristics are summarised in table 2.

MPI Subscales

Scores for the Pain Severity, Interference and Affective Distress MPI subscales were significantly higher in the poor outcome group across all time points ($p < 0.01$), and scores for the Life Control subscale were significantly lower in the poor outcome group across all time points ($p < 0.01$). The results of other scores are shown in Table 3.

In the good outcome group, the mean scores for Pain Severity, Interference, Support (all $p < 0.01$), and Affective Distress ($p = 0.02$) were significantly different over the study period with a downward trend as shown in figure 1.

In the poor outcome group, the Support subscale showed significantly different mean scores across the five study time points ($p < 0.01$) with a downward trend; the other subscales for this group did not differ as shown in figure 2.

Logistic regression

Spearman's correlation coefficients were calculated for all demographic variables, PHQ-4, GCPS score (0-4), and MPI subscale Rasch scores at each of the five time points. Four pairs of variables showed significant, strong, positive correlation across two or more time points (Spearman's $\rho > 0.6$, $p < 0.05$): Pain Severity subscale and GCPS (all time points), Interference subscale and GCPS (baseline, 24 months), Interference and Pain Severity subscales (baseline, 12, and 24 months), and Affective Distress subscale and PHQ-4 score (baseline, 12, 18, and 24 months). Two pairs of variables showed significant, strong, negative correlation across two or more time points (Spearman's $\rho < -0.6$, $p < 0.05$): Affective Distress and Life Control subscales (all time points) and Life Control Subscale and PHQ-4 (12 and 24 months). None of the variables showed correlation greater than $\rho = 0.76$, suggesting that multicollinearity is unlikely [27]. Correlation coefficients are presented in the Supplementary Data (online).

The only predictor in the logistic regression model whose odds ratio was significant ($p < 0.05$) was the Interference MPI subscale and this was the case across all time points. No other variables in the model had significant odds ratios at any time point. The odds ratio for Interference was between 1.14 and 1.98 across all time points and this is presented in table 4.

Discussion

Our findings indicate that people with POFP who have pain related disability which is consistent over time (i.e. those in the poor outcome group) are more likely to be younger, have higher IMD scores (i.e. are from more deprived areas), and a lower monthly income. They are more likely to have been in pain for longer, and consistently score higher on PHQ-4, indicating greater levels of psychosocial distress related to either anxiety and/or depression. It is impossible from our study design to discern whether more severe pain and disability is a result of these factors, or whether any causation is in the opposite direction. The study does, however, add support to previous work to suggest that there is a subgroup of patients with POFP whose needs are not met by current usual care and that standardised self-report questionnaires can be helpful in identifying this group. Our findings fit with current understanding of the aetiology and pathophysiology of these conditions and the impact of bio-psychosocial factors. The findings of the OPPERA study demonstrated that a number of psychological variables may be implicated in TMD development [6] and our results indicate that they continue to be associated with good or poor outcome throughout the course of POFP.

Pain Severity and Affective Distress scores were consistently higher in the poor outcome group indicating that the severity of participants' pain and the level of distress is associated with the presence of pain related disability over time (as reflected in the overall outcome based on dichotomised GCPS scores). Whilst it is perhaps unsurprising that the Interference subscale was consistently higher in the poor outcome group, as this subscale and the dichotomised GCPS both assess pain-related disability, this does indicate that the Interference subscale has good convergent validity with the GCPS. The Life Control subscale was consistently lower in the poor outcome group, perhaps indicating in these patients that disability is related to the degree of control they feel they have over their life and condition. Interestingly, we found that Support subscale scores were higher in the poor outcome group. This may reflect greater concern and attentiveness in the significant others of patients whose POFP has a greater impact on their function. This finding was not consistent over all time points however, and we also found that in the poor outcome group the Support subscale had a downward trend over the study period, despite the fact that the

other subscales in this group did not change over this time; this may indicate that perceived support from significant others diminishes over time.

In the good outcome group, Pain Severity, Interference, and Affective Distress all trended downward over the study period; this may indicate that for patients with a good disability-related outcome, it is not just measures of disability (i.e. the Interference subscale) that tend to improve over time, but also those relating to the severity of pain and to negative affect. The Support subscale also decreased over the study period in this group, indicating perhaps that as pain, distress, and disability improved, support displayed by significant others decreased also. At baseline, the scores for Distracting and Solicitous Responses subscales were similar between good and poor outcome groups, however the Punishing Responses subscale was higher in the poor outcome group. This indicates perhaps that patients' perception of the negative responses of their significant others to their pain is related to the severity of their condition.

The consistent downward trend which was present in interference, pain severity and affective distress indicates another way of distinguishing patients who are making good progress from those who might require a different or more intensive treatment approach. Based on our results we would recommend careful assessment and treatment review of patients who do not show an improvement within a 6-month period.

Comparison of MPI subscale scores to those of a sample of 6,532 heterogeneous chronic pain patients contained within the MPI version 3 software package (University of Pittsburgh, 2013) shows generally comparable scores in the poor outcome group, and more favourable scores in the good outcome group of the present study. Comparison of scores to other studies using the MPI in orofacial pain populations is difficult as previous investigators have generally reported raw MPI scores [14, 17, 18], or T-scores [17, 18, 28], instead of Rasch scores as we present. The Rasch model transforms the ordinal raw MPI scores and allows the resulting values to be treated as interval data [29]. One study reported a mean Rasch score for the Interference MPI subscale in patients with POFP in secondary care of 44.23 [16]; this is slightly higher than the scores reported for all participants in the present study, but falls between the scores we report for the good and poor outcome groups. Age and gender distribution were similar in this study, however the authors did not report the other MPI subscales.

Subgroup classifications have been derived (adaptive coper, interpersonally distressed, and dysfunctional clusters) by cluster analysis using MPI subscales and these are intended to reflect respondents' coping pattern to their pain [8]. These clusters have shown utility in a number of patient groups [8], including those with TMD [14, 30]. Although the test-retest stability of the individual MPI subscales is good, it has been suggested that these derived clusters demonstrate low stability [31]. Given that the MPI was administered on multiple occasions in the present study, and that the cluster analysis would not be easy to score routinely in clinical practice we elected not to use these clusters in our analyses.

Strong positive correlation (Spearman's $\rho > 0.6$) was seen across multiple time points between both the Pain Severity subscale and GCPS score (0 – 4) and the Interference subscale and GCPS score. This would suggest that both subscales have good convergent validity with the GCPS which is commonly used in this patient population. Additionally, both the Pain Severity and Interference subscales showed consistent strong correlation. This finding is consistent with that of other investigators who have demonstrated univariate association between other measures of pain intensity and pain related disability in diverse groups of chronic pain patients [32, 33]. Strong positive correlation was consistently seen between the Affective Distress subscale and PHQ-4 score. The Affective Distress subscale was also strongly, negatively correlated with the Life Control Subscale, which in turn was negatively correlated with PHQ-4 score. This may suggest the widespread use of the PHQ-4 in this patient group is likely to capture that which is captured using the Life Control and Affective Distress subscales of the MPI; this is important as the PHQ-4 is considerably easier to administer and score.

The logistic regression models indicated that the Interference subscale was the only consistent predictor of outcome group when other factors were accounted for (demographic variables, other MPI subscales, PHQ-4, and months in pain). Excluding the odds ratio at 12 months (which was 1.98), the odds ratio for the Interference subscale ranged from 1.14 to 1.26 depending on the time point; this suggests that a one point increase on the Interference subscale (Rasch score from 0 – 100, not raw score), equates to a 14 – 26% increase in the odds of being in the good outcome group. These findings are consistent with the recommendations of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT), who suggest that the MPI Interference subscale

should be one of two possible measures used to assess physical functioning in clinical trials for chronic pain (the other being the interference items of the Brief Pain Inventory) [34]. Our findings may not be surprising, given that the Interference subscale measures pain related disability, as does the dichotomised GCPS, however it is interesting that the other predictors do not emerge as significant in the regression model. This may indicate perhaps that in this patient group, pain related disability may be less strongly related to other factors such as pain severity; this is consistent with the findings of Karasawa *et al.* [33] who demonstrated that self-efficacy but not pain intensity predicted pain related disability in their multivariate analysis of patients in a chronic pain clinic. This is in contrast to non-chronic pain conditions such as acute back pain [35] where pain intensity predicts pain related disability, likely due to the lower impact of other factors in the acute stages of illness. Su *et al.* [36], in their study of patients with TMD, found that depression was a significant predictor of pain related disability using multivariate analysis. We found significantly higher PHQ-4 scores in the poor outcome group however PHQ-4 did not emerge as a significant predictor in our multivariate model. Su *et al.* did not include demographic variables such as deprivation, education and income in their regression model as we did, which may help to explain this difference.

Our study had a number of limitations which may limit the validity of our findings. First, we relied on dichotomised GCPS as our outcome measure, which essentially makes the instrument a measure of pain related disability [23]. We elected to use disability as our outcome as this reflects more of the impact of pain on patients' lives than, for example, pain intensity. We also chose to rely on mode dichotomised GCPS across all five time points, as this would better reflect the fluctuating nature of POFP over time. It is possible that our positive findings were because the Interference subscale also measures disability, however, the correlation coefficient for GCPS and Interference subscale scores were below that which would suggest the presence of multicollinearity. Additionally, our finding that other variables were not significant predictors in the model suggest that these variables may have more limited value in predicting outcome over time. It is true however that reliance on disability as an outcome measure may miss other features of patients' conditions, such as their ability to adapt to, and accept an altered level of functioning – and for some this may be a favourable outcome given the unremitting course of POFP in some patients.

Second, the number of participants in the poor outcome group was small (*n*=36) compared to the good outcome group (*n*=110), which will reduce the obtained statistical power compared to our *a priori* calculation. As the groups were derived at the end of the study based on the course of participant's conditions and not at the start, it was not possible to predict eventual group sizes prior to recruitment. Future studies using a similar outcome measure should consider this when recruiting participants. Similarly, the initial power calculation was performed to detect a difference between good and poor outcome groups in relation to GCPS and not for the logistic regression we performed; it is possible therefore that a similar study with greater power might detect other significant predictors in addition to our findings.

Third, although most variables had little missing data, some had a higher degree of missingness – income in particular had 25.3% missing data; income was lower in the poor outcome group, although it did not show significant correlation with any other variable in the unimputed data, and was not a significant predictor of outcome in the logistic regression models. It is possible that these findings may be different with complete data, although the findings for income were the same in logistic regression models using both imputed and non-imputed datasets (data not presented).

Fourth, because pain-related disability is a relatively common finding in patients with POFP (>10%) it is possible that the odds ratios obtained from logistic regression in this study overestimates the relative risk of pain-related disability. These odds ratios should therefore be interpreted in light of this. Further studies may consider alternative statistical methods to control for this (https://doi.org/10.1093/aje/kwg074).

Finally, the method used for imputation was MICE, and although this method does not have as robust an underlying mathematical basis as other methods such as Multivariate Normal imputation (MVN), MICE is widely used in the medical literature [37]. The flexibility of MICE allows imputation of multiple data types (categorical and continuous) and does not assume a multivariate normal distribution. Comparison of this method to complete case analysis shows reduced bias, and this is comparable to imputation using MVN [38]. Importantly, our key findings (i.e. Interference subscale was the only consistent, significant predictor of outcome) were identical in both imputed and non-imputed logistic regression models (data not presented).

Conclusions

We found that the MPI Interference subscale was a consistent predictor of clinical outcome based on pain-related disability in patients with POFP over time; this was true taking into account demographic variables, duration of pain, anxiety, and depression. The Interference and Pain Severity MPI subscales showed good convergent validity with the GCPS, as did the Life Control and Affective Distress subscales with the PHQ-4, suggesting that these briefer measures (GCPS, PHQ-4) may be more convenient for clinical use. Given this, the Interference subscale may be clinically useful in identifying patient with POFP likely to benefit from earlier or different clinical intervention . It is likely that the PHQ-4 covers the constructs measured in the Life Control and Affective Distress subscales in a more abbreviated measure.

Declarations

JRA, CP, JD contributed to study design and critically revised the manuscript, JRA and JD contributed to data analysis, JD contributed to initial study conception and data acquisition. All authors declare that there are no conflicts of interest, gave final approval, and agree to be accountable for all aspects of the work presented.

Data Availability

Reasonable requests for access to the anonymised participant level data will be considered by the authors.

References

1. International Classification of Orofacial Pain. 1st edition (ICOP). Cephalalgia. 2020;40(2):129-221.

2. Haggman-Henrikson B, Liv P, Ilgunas A, Visscher CM, Lobbezoo F, Durham J, et al. Increasing gender differences in the prevalence and chronification of orofacial pain in the population. Pain. 2020;DOI: 10.1097/j.pain.0000000000001872.

3. Forssell H, Sipila K, Teerijoki-Oksa T, Vartiainen P, Kautiainen H, Sintonen H, et al. The impact of chronic orofacial pain on health-related quality of life. Scand J Pain. 2019;DOI: 10.1515/sjpain-2019-0131.

4. Beecroft EV, Durham J, Thomson P. Retrospective examination of the healthcare 'journey' of chronic orofacial pain patients referred to oral and maxillofacial surgery. British Dental Journal. 2013;214(5):E12-E.

5. Macfarlane TV, Blinkhorn AS, Davies RM, Kincey J, Worthington HV. Predictors of Outcome for Orofacial Pain in the General Population: a Four-year Follow-up Study. Journal of Dental Research. 2004;83(9):712-7.

6. Fillingim RB, Ohrbach R, Greenspan JD, Knott C, Diatchenko L, Dubner R, et al. Psychological factors associated with development of TMD: the OPPERA prospective cohort study. The journal of pain : official journal of the American Pain Society. 2013;14(12 Suppl):T75-T90.

7. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). Pain. 1985;23(4):345-56.

8. Turk DC, Rudy TE. Toward an empirically derived taxonomy of chronic pain patients: integration of psychological assessment data. Journal of consulting and clinical psychology. 1988;56(2):233-8.

9. Andreu Y, Galdon MJ, Durá E, Ferrando M, Pascual J, Turk DC, et al. An examination of the psychometric structure of the Multidimensional Pain Inventory in temporomandibular disorder patients: a confirmatory factor analysis. Head & face medicine. 2006;2:48-.

10. Rudy TE. Multidimensional Pain Inventory Version 3.0 User's Guide. Pittsburg, PA: University of Pittsburgh; 2004.

11. Rudy TE, Turk DC, Zaki HS, Curtin HD. An empirical taxometric alternative to traditional classification of temporomandibular disorders. Pain. 1989;36(3):311-20.

12. Dahlström L, Widmark G, Carlsson SG. Cognitive-behavioral profiles among different categories of orofacial pain patients: diagnostic and treatment implications. European journal of oral sciences. 1997;105(5 Pt 1):377-83.

13. Reissmann DR, John MT, Wassell RW, Hinz A. Psychosocial profiles of diagnostic subgroups of temporomandibular disorder patients. European journal of oral sciences. 2008;116(3):237-44.

14. Turk DC, Rudy TE. The robustness of an empirically derived taxonomy of chronic pain patients. Pain. 1990;43(1):27-35.

15. Zucoloto ML, Maroco J, Duarte Bonini Campos JA. Psychometric Properties of the Multidimensional Pain Inventory Applied to Brazilian Patients with Orofacial Pain. Journal of oral & facial pain and headache. 2015;29(4):363-9.

16. Boggero IA, Rojas Ramirez MV, King CD. Cross-Sectional Associations of Fatigue Subtypes with Pain Interference in Younger, Middle-Aged, and Older Adults with Chronic Orofacial Pain. Pain medicine (Malden, Mass). 2020;DOI: 10.1093/pm/pnaa092.

17. Carlson CR, Miller CS, Reid KI. Psychosocial profiles of patients with burning mouth syndrome. Journal of orofacial pain. 2000;14(1):59-64.

18. Danhauer SC, Miller CS, Rhodus NL, Carlson CR. Impact of criteria-based diagnosis of burning mouth syndrome on treatment outcome. Journal of orofacial pain. 2002;16(4):305-11.

19. Durham J, Breckons M, Araujo-Soares V, Exley C, Steele J, Vale L. Developing Effective and Efficient care pathways in chronic Pain: DEEP study protocol. BMC oral health. 2014;14:6.

20. Gonzalez YM, Schiffman E, Gordon SM, Seago B, Truelove EL, Slade G, et al. Development of a brief and effective temporomandibular disorder pain screening questionnaire: reliability and validity. J Am Dent Assoc. 2011;142(10):1183-91.

21. Hapak L, Gordon A, Locker D, Shandling M, Mock D, Tenenbaum HC. Differentiation between musculoligamentous, dentoalveolar, and neurologically based craniofacial pain with a diagnostic questionnaire. Journal of orofacial pain. 1994;8(4):357-68.

22. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain. 1992;50(2):133-49.

23. Dworkin SF, Huggins KH, Wilson L, Mancl L, Turner J, Massoth D, et al. A randomized clinical trial using research diagnostic criteria for temporomandibular disorders-axis II to target clinic cases for a tailored self-care TMD treatment program. Journal of orofacial pain. 2002;16(1):48-63.

24. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ–4. Psychosomatics. 2009;50(6):613-21.

25. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. Stat Methods Med Res. 2007;16(3):219-42.

26. Breckons M, Shen J, Bunga J, Vale L, Durham J. DEEP Study: Indirect and Out-of-pocket Costs of Persistent Orofacial Pain. J Dent Res. 2018;97(11):1200-6.

27. Tabachnick BG. Using multivariate statistics 6th ed. Fidell LS, editor. Boston: Pearson Education; 2013.

28. Porto F, de Leeuw R, Evans DR, Carlson CR, Yepes JF, Branscum A, et al. Differences in psychosocial functioning and sleep quality between idiopathic continuous orofacial neuropathic pain patients and chronic masticatory muscle pain patients. Journal of orofacial pain. 2011;25(2):117-24.

29. Tennant A, McKenna SP, Hagell P. Application of Rasch Analysis in the Development and Application of Quality of Life Instruments. Value in Health. 2004;7:S22-S6.

30. Epker J, Gatchel RJ. Coping profile differences in the biopsychosocial functioning of patients with temporomandibular disorder. Psychosom Med. 2000;62(1):69-75.

31. McKillop JM, Nielson WR. Improving the usefulness of the Multidimensional Pain Inventory. Pain research & management. 2011;16(4):239-44.

32. Garbi Mde O, Hortense P, Gomez RR, da Silva Tde C, Castanho AC, Sousa FA. Pain intensity, disability and depression in individuals with chronic back pain. Rev Lat Am Enfermagem. 2014;22(4):569-75.

33. Karasawa Y, Yamada K, Iseki M, Yamaguchi M, Murakami Y, Tamagawa T, et al. Association between change in self-efficacy and reduction in disability among patients with chronic pain. PLOS ONE. 2019;14(4):e0215404.

34. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005;113(1).

35. Salt E, Wiggins AT, Hooker Q, Crofford L, Rayens MK, Segerstrom S. The Effects of Pain Severity, Pain Catastrophizing, Depression, and Exercise on Perceived Disability in Acute Low Back Pain Patients. Res Theory Nurs Pract. 2018;32(4):436-48.

36. Su N, Lobbezoo F, van Wijk A, van der Heijden GJMG, Visscher CM. Associations of pain intensity and pain-related disability with psychological and socio-demographic factors in patients with temporomandibular disorders: a cross-sectional study at a specialised dental clinic. Journal of Oral Rehabilitation. 2017;44(3):187-96.

37. Hayati Rezvan P, Lee KJ, Simpson JA. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC Medical Research Methodology. 2015;15(1):30.

38. Lee KJ, Carlin JB. Multiple Imputation for Missing Data: Fully Conditional Specification Versus Multivariate Normal Imputation. American Journal of Epidemiology. 2010;171(5):624-32.

Tables

	Missing observations	*% missing*
Overall GCPS Outcome	0	0.0
Age	0	0.0
Gender	0	0.0
Months in pain	0	0.0
IMD score	2	1.4
Education classification	14	9.6
Monthly income	37	25.3
PHQ-4 (baseline)	7	4.8
PHQ-4 (12 months)	4	2.7
PHQ-4 (24 months)	29	19.9
MPI (baseline)	1	0.7
MPI (6 months)	2	1.4
MPI (12 months)	1	0.7
MPI (18 months)	13	8.9
MPI (24 months)	16	11.0

Table 1. Missing observations for each variable. GCPS: Graded Chronic Pain Scale, IMD: Index of Multiple Deprivation, PHQ-4: Patient Health Questionnaire-4, MPI: Multidimensional Pain Inventory.

	All Participants	***Good Outcome Group***	***Poor Outcome Group***
n	146	110	36
Mean age, years (*SD*)	53.25 (14.6)	55.26 (14.4)	47.11 (13.5) **
Gender, % female	82.19	82.73	80.56
Education, % with degree	43.94	42.42	48.48
Mean IMD score, (*SD*)	17.54 (12.7)	16.07 (12.4)	21.96 (12.7) **
Mean monthly income, £ (*SD*)	1203.72 (935.6)	1315.90 (972.3)	908.32 (769.9) *
Mean months in pain, (*SD*)	104.25 (117.0)	98.09 (118.7)	123.06 (111.2) *
Median PHQ-4 [baseline], (*IQR*)	2 (4)	2 (4)	4.5 (6) **
Median PHQ-4 [12 months], (*IQR*)	2 (4)	1 (3)	4.5 (7) **
Median PHQ-4 [24 months], (*IQR*)	2 (5)	1.5 (3)	6 (7) **

Table 2. Demographic characteristics for all participants and by overall GCPS outcome. SD: Standard Deviation. IMD: Index of Multiple Deprivation. PHQ-4: Patient Health Questionnaire-4. IQR: Interquartile Range. Comparison between good and poor outcome groups: *$p < 0.05$, **$p < 0.01$.

MPI Subscale Rasch Score		*Good Outcome Group*	*Poor Outcome Group*
Baseline	**Pain Severity**	39.4 (16.9)	57.5 (13.8) **
n = 145	**Interference**	36.9 (14.7)	54.1 (8.7) **
	Life Control	66.2 (21.9)	47.9 (11.0) **
	Affective Distress	44.1 (15.1)	59.4 (14.1) **
	Support	48.9 (28.1)	60.8 (25.6) *
	Punishing Responses	22.9 (27.2)	41.6 (29.9) **
	Solicitous Responses	51.2 (18.3)	58.6 (20.9)
	Distracting Responses	45.9 (26.6)	49.3 (28.0)
6m	**Pain Severity**	36.8 (20.1)	55.3 (16.1) **
n = 144	**Interference**	35.2 (14.9)	54.3 (8.0) **
	Life Control	66.0 (19.5)	47.4 (13.0) **
	Affective Distress	42.8 (15.0)	57.2 (14.7) **
	Support	46.2 (26.3)	57.0 (26.6) *
12m	**Pain Severity**	31.6 (16.8)	56.5 (17.4) **
n = 145	**Interference**	31.1 (14.1)	54.5 (8.9) **
	Life Control	66.1 (22.1)	47.0 (15.6) **
	Affective Distress	41.7 (15.9)	59.1 (13.7) **
	Support	43.0 (27.8)	59.4 (23.8) **
18m	**Pain Severity**	33.5 (18.1)	52.6 (13.8) **
n = 133	**Interference**	31.0 (16.5)	54.3 (10.2) **
	Life Control	68.4 (19.7)	48.3 (12.1) **
	Affective Distress	41.3 (17.0)	55.0 (13.1) **
	Support	44.0 (30.2)	54.8 (25.4)
24m	**Pain Severity**	30.3 (19.2)	52.5 (18.9) **
n = 130	**Interference**	27.5 (16.1)	55.2 (13.4) **
	Life Control	66.6 (21.5)	46.1 (15.7) **
	Affective Distress	40.2 (16.9)	58.4 (18.3) **
	Support	41.7 (29.6)	50.2 (28.5)

Table 3. Mean (standard deviation) Multidimensional Pain Inventory (MPI) subscale Rasch scores at each time point by overall GCPS outcome group. Comparison between good and poor outcome groups: Wilcoxon rank-sum, *$p < 0.05$, **$p < 0.01$.

Time point	*MPI Subscale*	*Predictor odds ratio (95% CI)*	*p*
Baseline	Pain Severity	1.01 (0.97-1.05)	0.672
	Interference	1.14 (1.05-1.24)	0.003*
	Life Control	0.99 (0.94-1.04)	0.632
	Affective Distress	1.02 (0.94-1.10)	0.641
	Support	1.00 (0.98-1.03)	0.626
	Punishing Responses	1.01 (0.99-1.03)	0.159
	Solicitous Responses	1.02 (0.97-1.08)	0.351
	Distracting Responses	0.98 (0.96-1.01)	0.151
6m	Pain Severity	0.99 (0.95-1.04)	0.749
	Interference	1.26 (1.12-1.41)	0.000*
	Life Control	0.97 (0.91-1.03)	0.269
	Affective Distress	1.01 (0.95-1.07)	0.777
	Support	1.00 (0.98-1.03)	0.698
12m	Pain Severity	0.93 (0.83-1.03)	0.164
	Interference	1.98 (1.06-3.69)	0.032*
	Life Control	0.99 (0.89-1.11)	0.907
	Affective Distress	1.06 (0.90-1.24)	0.493
	Support	1.03 (0.99-1.07)	0.191
18m	Pain Severity	0.99 (0.95-1.05)	0.977
	Interference	1.17 (1.05-1.30)	0.004*
	Life Control	0.95 (0.89-1.01)	0.118
	Affective Distress	0.99 (0.94-1.05)	0.834
	Support	1.00 (0.97-1.02)	0.703
24m	Pain Severity	0.97 (0.91-1.04)	0.374
	Interference	1.21 (1.04-1.40)	0.013*
	Life Control	0.99 (0.93-1.04)	0.633
	Affective Distress	1.02 (0.96-1.08)	0.565
	Support	1.00 (0.97-1.03)	0.870

Table 4. Outcome of logistic regression with overall GCPS outcome as dependant variable at baseline, 6, 12, 18, and 24 months. Multidimensional Pain Inventory (MPI) subscale Rasch scores were predictors of interest reported here, with demographic variables (age, gender, education, deprivation, months in pain, income) and PHQ-4 scores also included as additional predictors. This analysis used multiple imputed datasets (m=10) generated using chained equations. 95%CI: 95% confidence interval. *$p < 0.05$.

Figures

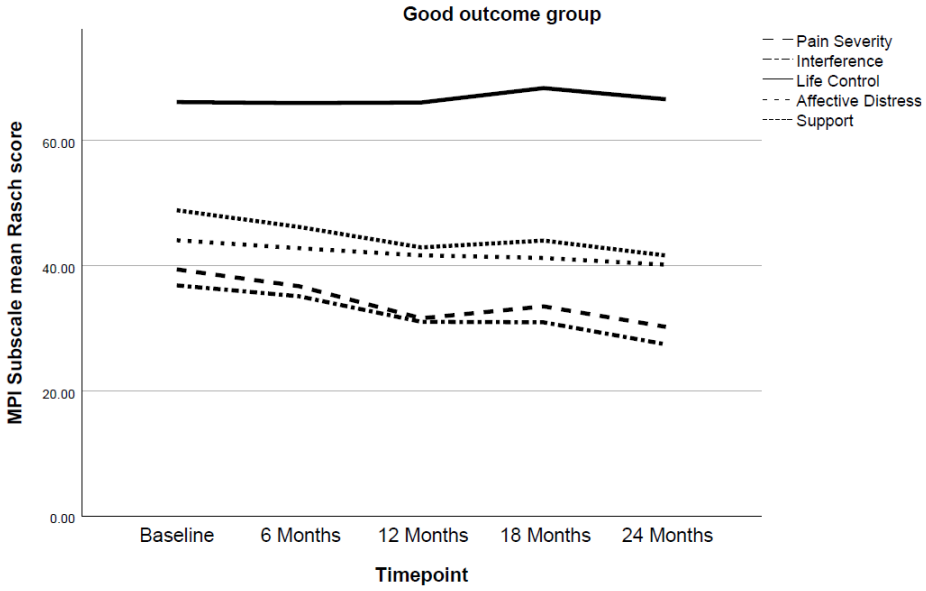

Figure 1. Line **c**hart showing trends in mean Multidimensional Pain Inventory (MPI) subscale Rasch scores in the good outcome group. Significant differences were seen across time points for the Pain Severity, Interference, Support (Friedman test, all $p < 0.01$), and Affective Distress ($p = 0.02$) scores. There was no difference in scores of the Life Control subscale ($p = 0.60$). Standard deviation is not shown on the chart to aid readability; however, this is presented in Table 3.

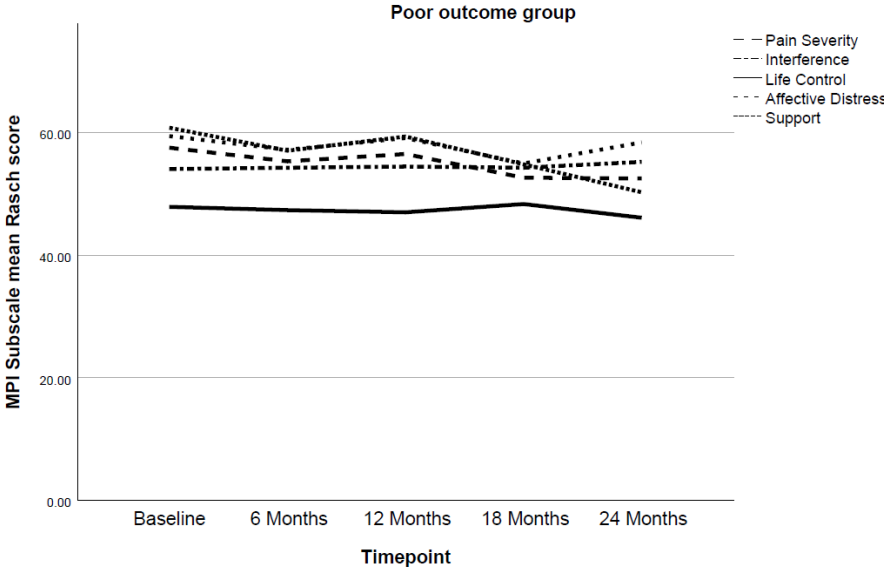

Figure 2. Line chart showing trends in mean Multidimensional Pain Inventory (MPI) subscale Rasch scores in the poor outcome group. Significant differences were seen across time points for the Support MPI subscale (Friedman test, $p < 0.01$), but not for other subscales (all $p > 0.05$). Standard deviation is not shown on the chart to aid readability; however, this is presented in Table 3.

DEEP Study: Utility of the multidimensional pain inventory in persistent orofacial pain—Supplementary Data

James R Allison[1, 2 *], Chris Penlington[1, 2], Justin Durham[1, 2]

Correlation coefficients

Recruitment

	GCPS	Age	Gender	Education	IMD	Months in Pain	Income	PHQ-4	Pain Severity	Interference	Life Control	Affective Distress	Support
GCPS	1.0000												
Age	-0.2219*	1.0000											
Gender	0.0517	0.0696	1.0000										
Education	0.0053	0.2343*	-0.0105	1.0000									
IMD	0.2524*	-0.2800*	0.0978	0.1795	1.0000								
Months in Pain	0.0563	-0.0004	0.1452	-0.1520	0.1059	1.0000							
Income	-0.1170	-0.0894	0.1279	-0.2900*	-0.1773	-0.0540	1.0000						
PHQ-4	0.3731*	-0.2551*	0.0304	-0.1174	0.1569	-0.0134	-0.0919	1.0000					
Pain Severity	0.6086*	-0.1487	0.0355	0.0320	0.2093*	0.1325	-0.2575*	0.3273*	1.0000				
Interference	0.6766*	-0.2680*	-0.0306	-0.1508	0.1774	0.2501*	-0.1500	0.3771*	0.6279*	1.0000			
Life Control	-0.4211*	0.4388*	-0.0193	0.2866*	-0.1941	-0.1039	0.1022	-0.5559*	-0.4709*	-0.4964*	1.0000		
Affective Distress	0.3981*	-0.2576*	0.0621	-0.1155	0.2435*	0.0759	-0.0823	0.6010*	0.5447*	0.4761*	-0.6962*	1.0000	
Support	0.2164*	0.0316	0.1404	0.1085	0.0292	0.2233*	-0.0542	-0.1847	0.1279	0.2123*	0.0378	0.0340	1.0000

Table S1. Spearman correlation coefficients for all variables at baseline. *: p<0.05

6 Months

	GCPS	Age	Gender	Education	IMD	Months in Pain	Income	PHQ-4	Pain Severity	Interference	Life Control	Affective Distress	Support
GCPS	1.0000												
Age	0.0155	1.0000											
Gender	-0.0652	0.0955	1.0000										
Education	0.0603	0.2344*	-0.0035	1.0000									
IMD	0.2628*	-0.2772*	0.0932	0.1783	1.0000								
Months in Pain	0.1881	0.0132	0.1739	-0.1460	0.1063	1.0000							
Income	-0.2272*	-0.0859	0.1303	-0.2896*	-0.1797	-0.0515	1.0000						
PHQ-4	0.1196	-0.2590*	0.0133	-0.1186	0.1576	-0.0197	-0.0935	1.0000					
Pain Severity	0.6700*	0.0087	0.0741	0.1944	0.1515	0.1379	-0.3092*	0.0727	1.0000				
Interference	0.4858*	-0.2032*	-0.0563	-0.1524	0.0993	0.2458*	-0.2062*	0.3197*	0.5662*	1.0000			
Life Control	-0.2472*	0.1586	0.0838	0.0978	-0.0467	0.0034	0.2026*	-0.4800*	-0.2965*	-0.4668*	1.0000		
Affective Distress	0.3645*	-0.2260*	0.0501	0.0059	0.1548	-0.0058	-0.1978	0.3630*	0.4903*	0.4171*	-0.6102*	1.0000	
Support	0.1606	0.0444	0.0775	0.1102	-0.0063	0.1120	-0.0651	-0.0809	0.1933	0.2430*	0.0962	0.0691	1.0000

Table S2. Spearman correlation coefficients for all variables. MPI subscales at 6 months, all other variables at baseline. *: p<0.05

12 Months

	GCPS	Age	Gender	Education	IMD	Months in Pain	Income	PHQ-4	Pain Severity	Interference	Life Control	Affective Distress	Support
GCPS	1.0000												
Age	0.0220	1.0000											
Gender	-0.0078	0.1091	1.0000										
Education	0.0174	0.2100*	0.0143	1.0000									
IMD	0.2204*	-0.2786*	0.0945	0.1691	1.0000								
Months in Pain	0.3432*	0.0088	0.1694	-0.1694	0.0859	1.0000							
Income	-0.2339*	-0.0953	0.1164	-0.2630*	-0.1583	-0.0435	1.0000						
PHQ-4 (12 months)	0.2305*	-0.3186*	0.0459	-0.1201	0.1038	-0.0355	-0.0300	1.0000					
Pain Severity	0.7085*	-0.0800	-0.0090	0.1030	0.2180*	0.2390*	-0.2394*	0.3048*	1.0000				
Interference	0.5784*	-0.1775	-0.0034	-0.1118	0.1772	0.2645*	-0.2361*	0.4227*	0.6530*	1.0000			
Life Control	-0.3698*	0.3236*	-0.0017	0.1386	-0.1713	-0.0257	-0.0313	-0.6045*	-0.3611*	-0.4391*	1.0000		
Affective Distress	0.3519*	-0.3287*	0.0387	-0.0558	0.2292*	-0.0924	-0.0429	0.7324*	0.5337*	0.4794*	-0.6411*	1.0000	
Support	0.1884	0.1249	-0.0292	0.2168*	0.0322	0.0621	-0.1714	-0.1118	0.1905	0.2610*	0.1441	0.0223	1.0000

Table S3. Spearman correlation coefficients for all variables. MPI subscales and PHQ-4 at 12 months, all other variables at baseline. *: $p<0.05$

18 Months

	GCPS	Age	Gender	Education	IMD	Months in Pain	Income	PHQ-4	Pain Severity	Interference	Life Control	Affective Distress	Support
GCPS	1.0000												
Age	-0.1853	1.0000											
Gender	-0.1355	0.1249	1.0000										
Education	0.0247	0.2316*	0.0204	1.0000									
IMD	0.3519*	-0.2530*	0.0788	0.1216	1.0000								
Months in Pain	0.1104	0.0018	0.1742	-0.1447	0.1214	1.0000							
Income	-0.0950	-0.1119	0.1314	-0.2371*	-0.1260	-0.0514	1.0000						
PHQ-4 (12 Months)	0.3807*	-0.3350*	0.0341	-0.1072	0.1195	-0.0867	-0.0437	1.0000					
Pain Severity	0.6855*	-0.0391	0.0039	0.1655	0.2306*	0.2159*	-0.1855	0.2650*	1.0000				
Interference	0.5398*	-0.1746	0.0250	-0.0516	0.1567	0.1252	-0.0971	0.3802*	0.5301*	1.0000			
Life Control	-0.3777*	0.3350*	-0.0161	0.1945	-0.0902	0.0136	-0.1312	-0.5936*	-0.1754	-0.4066*	1.0000		
Affective Distress	0.3954*	-0.1589	0.0720	0.0260	0.1584	-0.1199	0.0243	0.6208*	0.3452*	0.3944*	-0.6554*	1.0000	
Support	0.1943	0.0860	0.0308	0.1875	0.0257	0.1621	-0.0504	-0.0445	0.2276*	0.2594*	0.0180	0.0011	1.0000

Table S4. Spearman correlation coefficients for all variables. MPI subscales at 18 months, PHQ-4 at 12 months, all other variables at baseline. *: $p<0.05$

24 Months

	GCPS	Age	Gender	Education	IMD	Months in Pain	Income	PHQ-4	Pain Severity	Interference	Life Control	Affective Distress	Support
GCPS	1.0000												
Age	-0.1461	1.0000											
Gender	0.0547	0.1271	1.0000										
Education	0.0364	0.2551*	-0.0054	1.0000									
IMD	0.1136	-0.2306*	0.1058	0.0991	1.0000								
Months in Pain	0.2181*	0.0116	0.2328*	-0.1416	0.0936	1.0000							
Income	-0.1962	-0.2000	0.1204	-0.2615*	-0.0495	-0.0809	1.0000						
PHQ-4 (24 Months)	0.5473*	-0.2833*	0.1384	0.0528	0.1737	-0.0102	-0.0159	1.0000					
Pain Severity	0.7648*	-0.0351	0.0538	0.1889	0.0902	0.2083	-0.2979*	0.5096*	1.0000				
Interference	0.6608*	-0.1823	0.0204	-0.0328	0.2115	0.2718*	-0.2568*	0.5021*	0.7130*	1.0000			
Life Control	-0.3642*	0.2678*	-0.0516	0.0283	-0.1805	-0.0206	0.0774	-0.6525*	-0.3039*	-0.4853*	1.0000		
Affective Distress	0.5042*	-0.1893	0.1001	0.0173	0.0879	-0.1430	-0.0665	0.7072*	0.4996*	0.4802*	-0.6683*	1.0000	
Support	0.2920*	0.1104	0.0762	0.1668	-0.0856	0.1826	-0.0575	-0.0037	0.3120*	0.1945	0.2245*	-0.0410	1.0000

Table S5. Spearman correlation coefficients for all variables. MPI subscales and PHQ-4 at 24 months, all other variables at baseline. *: $p<0.05$

Graded Chronic Pain Scale Distribution

Recruitment

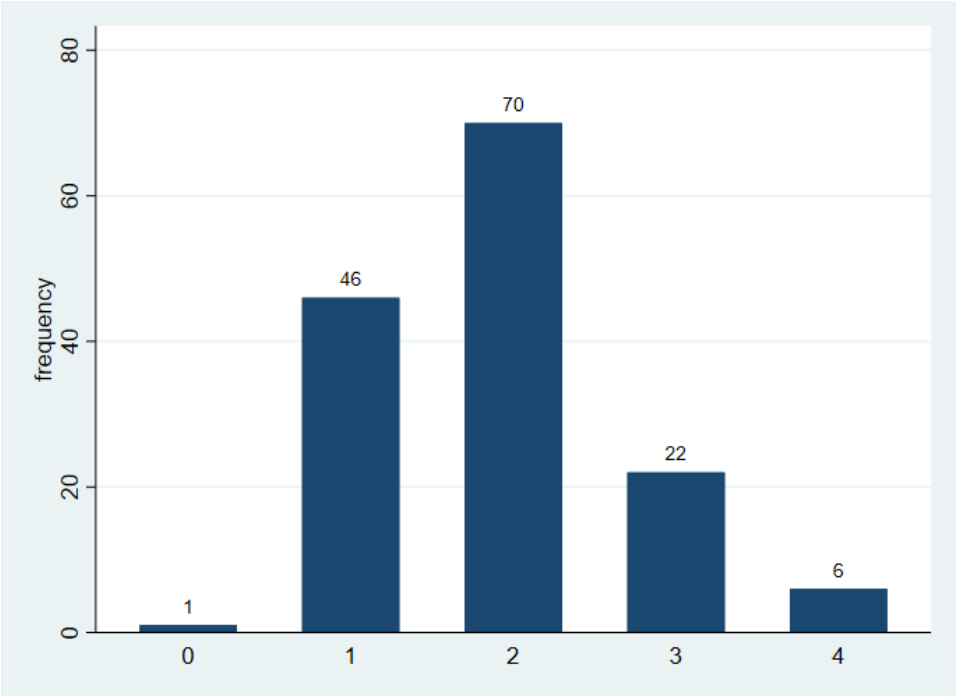

Figure S1. Distribution of Graded Chronic Pain Scale (GCPS) score for all participants at recruitment.

6 Months

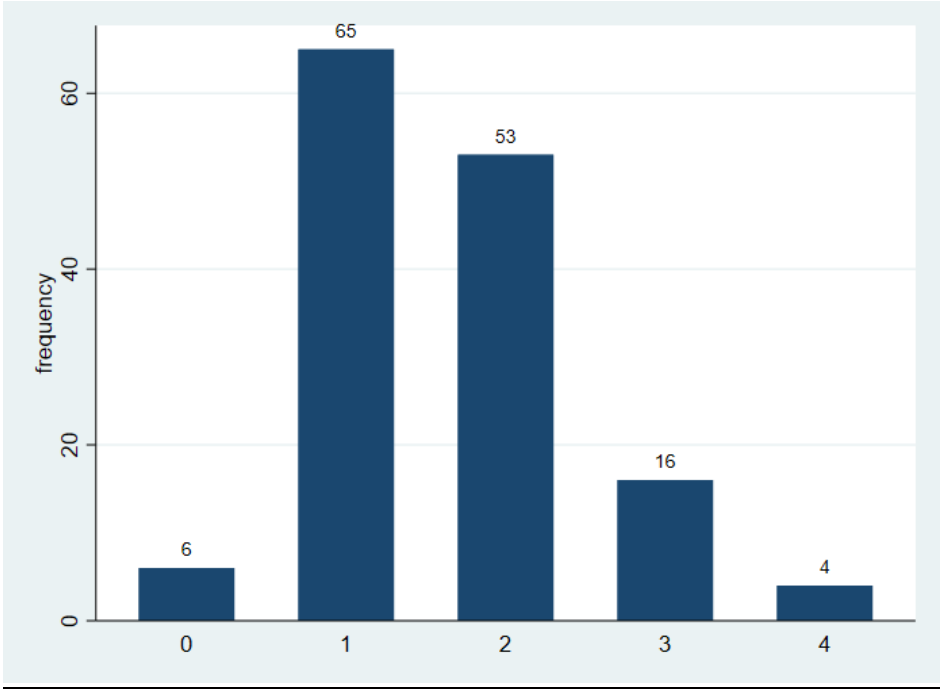

Figure S2. Distribution of Graded Chronic Pain Scale (GCPS) score for all participants at 6 months.

12 Months

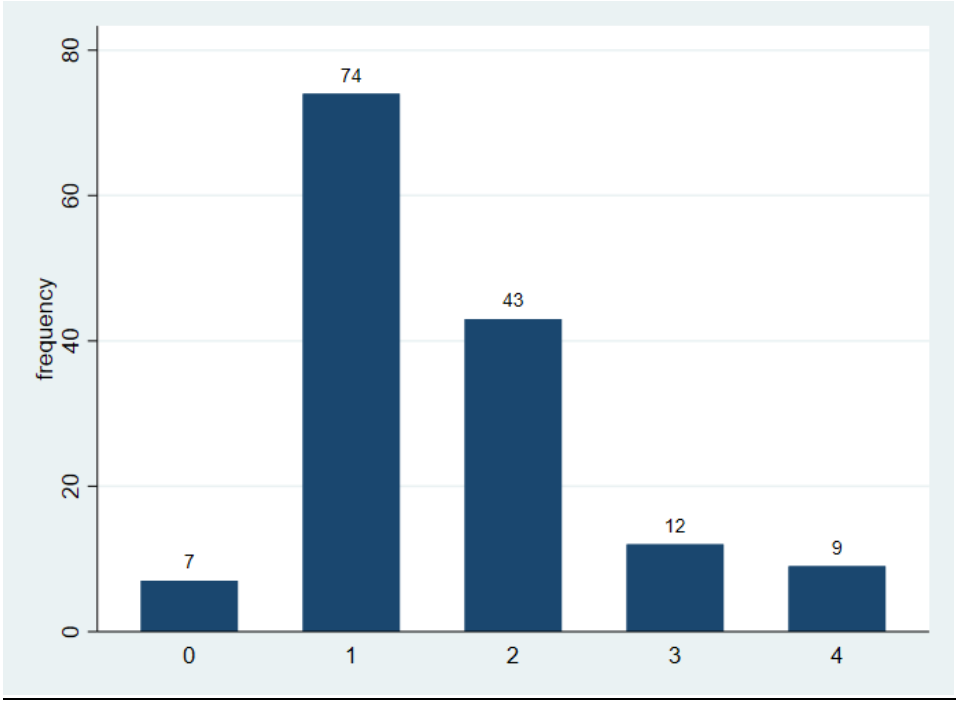

Figure S3. Distribution of Graded Chronic Pain Scale (GCPS) score for all participants at 12 months.

18 Months

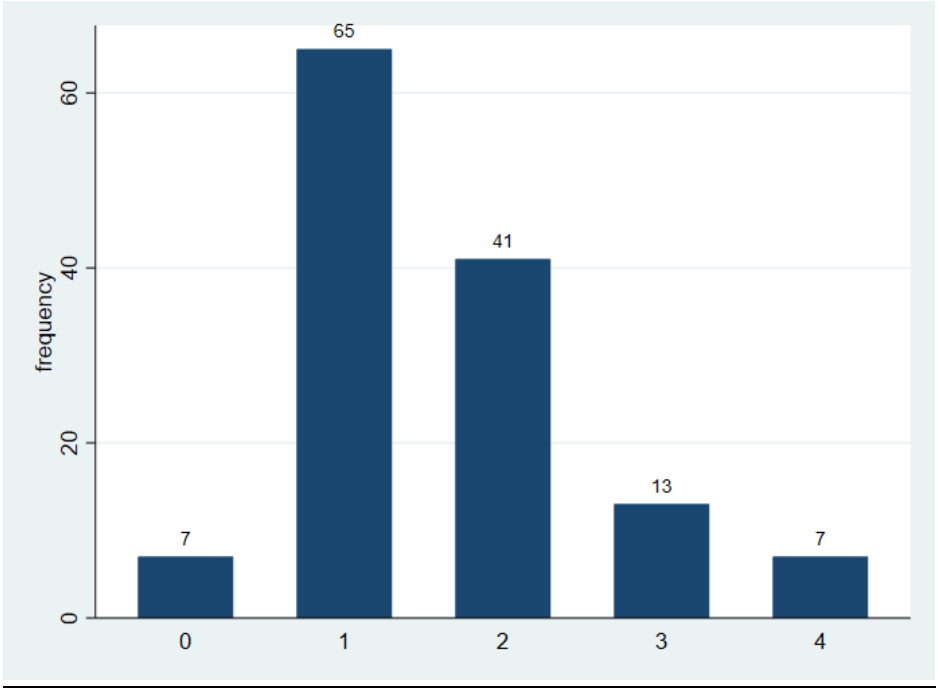

Figure S4. Distribution of Graded Chronic Pain Scale (GCPS) score for all participants at 18 months.

24 Months

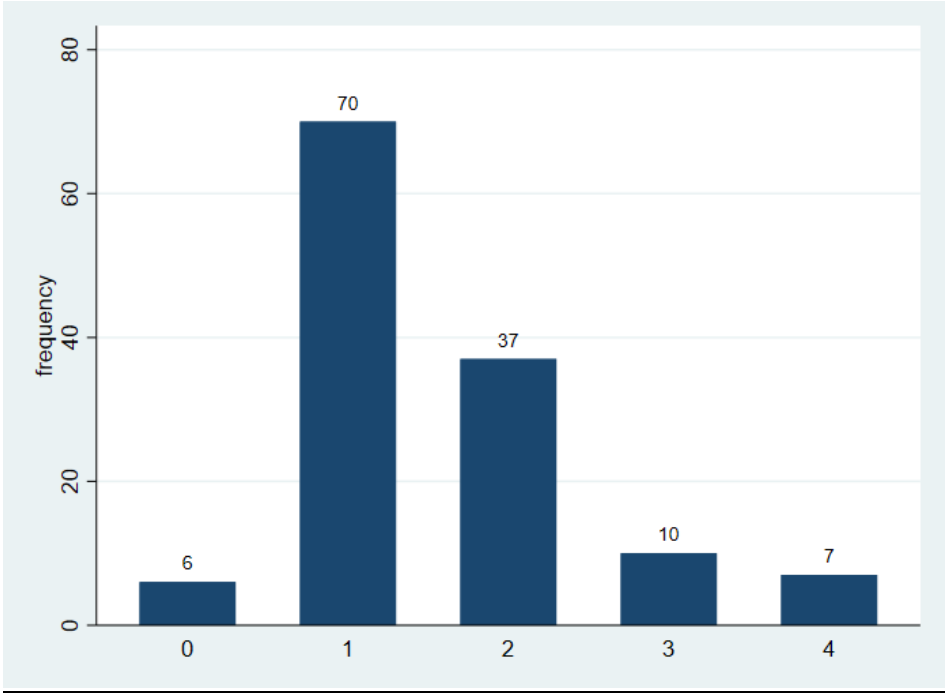

Figure S5. Distribution of Graded Chronic Pain Scale (GCPS) score for all participants at 24 months.