

Predicting Persistent Orofacial Pain: the role of illness perceptions, anxiety and depression.

Abstract Word Count 275

Total Word Count (Abstract to Acknowledgements) 3999

Total number of tables 4 (plus 8 in appendix of supplementary material)

Total number of figures 1

Number of references 34

Key words: chronic pain, facial pain, longitudinal studies, observational study, prognosis, depression.

Dr. Chris Penlington
Centre for Oral Health Research,
School of Dental Sciences
Newcastle University
Framlington Place
Newcastle
NE2 4BW

chris.penlington@ncl.ac.uk

Dr. Vera Araujo-Soares
Institute of Health and Society
Newcastle University
Baddiley-Clark Building
Richardson Road
Newcastle
NE2 4AX

vera.araujo-soares@ncl.ac.uk

Prof. Justin Durham
Centre for Oral Health Research,
School of Dental Sciences
Newcastle University
Framlington Place
Newcastle
NE2 4BW

justin.durham@ncl.ac.uk

Knowledge Transfer Statement

The results of this study are relevant because they will allow clinicians to consider the use of brief and easy-to-administer self-report measures to identify POFP patients at higher risk of poor outcome so that management can be planned accordingly.

Abstract

INTRODUCTION: Persistent orofacial pain (POFP) can be caused by a range of conditions affecting the mouth and face and is often associated with significant disability.

Biopsychosocial factors are known to be important predictors and have not yet been fully explored in this population.

OBJECTIVES: This study aimed to explore whether Illness perceptions and psychological distress (anxiety and depression) could predict long-term outcome in a community based sample of patients receiving routine care for persistent orofacial pain (POFP).

METHODS: A longitudinal cohort design study assessed 198 patients recruited from primary or secondary healthcare settings on 5 separate occasions over a 2 year period. Outcome was measured by the Graded Chronic Pain Scale and dichotomised into good or poor outcome categories. Independent variables included subscale scores of the psychometrically shortened Illness Perception Questionnaire (IPQ-PR) and the Patient Health Questionnaire-4 (PHQ-4, assessing anxiety and depressive symptomatology). Logistic regressions were performed to test whether scores on each subscale would be associated with different outcome for patients at 12 and 24 months and overall.

RESULTS: Beliefs about Consequences predicted outcome category (good vs poor) above and beyond other illness perception subscales, anxiety and depressive symptomatology ($z=3.78$, $p<0.000$, OR 3.05 95%CI 1.71 – 5.43). Both depressive symptomatology, measured by the PHQ-2 ($p=.001$ OR 4.06 95%CI 1.74 – 9.52) and psychological distress (mixed anxiety and depression, $p=.029$ OR 2.88 95%CI 1.12 – 7.41) measured by the PHQ-4 were also predictive of poor outcome, however these effects were no longer significant once measures of Illness perceptions were added.

CONCLUSION: Beliefs about the consequences of POFP are important predictors of outcome independent of other variables and can easily and briefly be included in assessments to inform management decisions.

Introduction

Persistent Orofacial Pain (POFP) refers to a range of pain conditions of a non-dental aetiology in the region of the face and mouth (Beecroft et al. 2013). POFP can be difficult to diagnose and manage and may have a considerable impact on quality of life (Shueb et al. 2015).

POFP is known to be biopsychosocial in nature (Dworkin 1994), influenced by a broad range of both psychosocial and physiological factors. Increased somatization (now classified by DSM-5 as somatic symptom disorder (APA 2013)), depression (Liao et al. 2011; Velly et al. 2011; Nevalainen et al. 2017) and health anxiety (Aggarwal et al. 2010) have been found to increase the risk of developing POFP and to predict poor outcome once pain is present (Galli et al. 2010; Velly et al. 2011). These findings mirror those in a wide range of persistent pain (Pincus et al. 2002) and other long-term conditions which are the basis of English government recommendations for psychological treatments for long-term conditions (National Collaborating Centre for Mental Health 2018).

An important question that remains is how we might draw on our knowledge of the relationships between these factors to develop, target and deliver appropriate management strategies. Referral to psychology services is recommended for patients with comorbid depression or anxiety (National Collaborating Centre for Mental Health 2018) but it can be difficult to collect relevant information to aid referral decisions within the short space of time available clinically. It is also difficult to use diagnoses of depression, anxiety or somatic symptom disorder to plan specific management strategies as such diagnoses cover a broad range of symptoms. While they provide information about a possible poor prognosis, a detailed understanding of mechanisms behind this effect may enable more effective targeting of treatment strategies.

Specific psychological constructs that may be associated with poor outcome in POFP conditions include catastrophization, self-efficacy and somatization (Velly et al. 2011) (Litt et al. 2010; Litt and Porto 2013). Illness representations have also been found to be important. A study (Galli et al. 2010) identified beliefs about the consequences of pain as an important predictor of outcome after 3 months of treatment, with beliefs about low personal control and a chronic timeline also accounting for some of the variation in outcomes. This is important because when illness beliefs can be readily addressed there is evidence that even a single session may lead to improved outcomes (Bonathan et al. 2014). Galli et al, however, reported outcomes only up to 6 months and within the context of a specialist Orofacial Pain Service where participants received individualised therapy according to their assessed needs. It is not known whether beliefs about illness at baseline would also be relevant within a community sample receiving non-specialist treatment over a longer period of time.

The primary aim of this study was to examine whether illness beliefs would predict outcome in a POFP cohort receiving routine treatment in the North of England over a period of 2 years. Secondary aims were to explore the ability of the Patient Health Questionnaire-4 (PHQ-4) which measures anxiety and depression and the Patient Health Questionnaire-2 (PHQ-2) which measures depression to screen for psychological comorbidity which may warrant further exploration and referral to specialist services. This was part of a larger study which tracked patient care pathways and costs associated with POFP over a 2 year period (Durham et al. 2016).

Methods

A closed cohort design longitudinal study was used to follow 198 patients receiving routine care for POFP in primary or secondary healthcare settings. A series of logistic regressions

were carried out to assess the relative contribution of attributes measured by a number of self-report measures to long-term outcome of POFP.

Methodology used in this study is consistent with STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for a human observational study. The study providing the data for this manuscript was reviewed and approved by a U.K. research ethics committee (NRES Reference: 12/YH/0338).

Sample and data collection

The sample consisted of patients who had experienced musculoskeletal, neuropathic, or neurovascular POFP for 3 months or more. Ethical guidelines were followed and all patients gave written, informed consent to participate in the study.

Full details of recruitment and data collection procedures are described elsewhere (Durham et al. 2014; Durham et al. 2016; Breckons et al. 2018) and are available with open access at eprint.ncl.ac.uk. Further information is also included in the supplementary appendix. Briefly, a power calculation was carried out based on an empirical decision aimed at ensuring it was possible to detect a difference in the Graded Chronic Pain Scale between those receiving treatment in primary and those in secondary care at a moderate effect size of 0.4 (Cohen 1992) and type 1 and 2 errors of 5% and 20% respectively ($\alpha=0.05$, $\beta=0.2$). The calculation indicated a total sample size of 200 would be required. This sample size also served the need to be able to identify significant predictors of pain-related disability according to accepted practice (Green 1991). The initial aim was to recruit 240 participants in order to allow for an estimated dropout rate of 20%. Data for this study consisted of the responses of 198 patients who completed baseline measures.

Primary Outcome Measure

The primary outcome measure for this study was the Graded Chronic Pain Scale (GCPS) (Von Korff et al. 1992). The GCPS is a measure of characteristic pain and disability widely used in this population. Pain intensity and disability are each measured by taking an average of three questions with scores ranging from 0 to 10. These scores are then combined with the number of days patients report being prevented from their usual activities in the last six months. At each of the five data collection timepoints (baseline, 6 months, 12 months, 18 months and 24 months) GCPS scores were dichotomised into 'low' GCPS (GCPS grades 0, 1 or 2a) or 'high' GCPS (GCPS grades 2b, 3 or 4) states following the method described by Dworkin and colleagues (Dworkin et al. 2002). Based on the rationale that POFP are long-term conditions which tend to vary over time, outcomes were further collapsed into a single outcome measure ('overall GCPS outcome over time'), according to the mode GCPS state over time for each participant. Thus, each participant with complete GCPS data was allocated to one of two groups: overall GCPS 'good outcome over time' (low GCPS state on at least 3 of the 5 study timepoints), or overall GCPS 'poor outcome over time' (high GCPS state on at least 3 of the 5 study timepoints).

Other Measures

The PHQ-4, an ultra-brief screening tool (Kroenke et al. 2009) screens for depression and anxiety by combining the PHQ-2 (for depression) with the Generalized Anxiety Disorder scale (GAD-2), an ultra-brief anxiety screener (Kroenke et al. 2007). Each scale includes two questions with responses on a four point scale ranging from 'not at all' to 'nearly every day'. The scales show good reliability, with Cronbach's alpha for this sample of 0.89 (PHQ-4) and 0.89 (PHQ-2). At each point in time, depression was coded as present (PHQ-2 score of 3 or above) or absent (PHQ-2 score of 2 or below). For comparison, the total PHQ-4 score was also calculated and collapsed into two categories (PHQ-4 score of 6 or above indicating

psychological distress). This corresponds to the recommended cut off point for detecting moderate depression and anxiety (Löwe et al. 2010).

The Illness Perceptions Questionnaire-Psychometrically Shortened (IPQ-PS) (Snihotta et al. 2010) is a shortened version of the Illness Perceptions Questionnaire-Revised (IPQ-R) (Moss-Morris et al. 2002) developed by using the three items for each IPQ-R subscale which demonstrated the highest factor loadings in the original study (Moss-Morris et al. 2002). The scale consists of 21 items across 7 domains (listed with Cronbach's alpha for this sample): Timeline (0.88), Consequences (0.82), Personal Control (0.60), Treatment Control (0.75), Illness Coherence (0.47) and Emotional Representations (0.84). A list of questions included is shown in Table S1. Each question is scored on a 5 point scale ranging from strongly disagree to strongly agree. Means and standard deviations of the independent variables used in the study at baseline are summarised in Table 2.

Data Analysis

A standardised protocol was followed for nonresponse which is reported in the supplementary appendix. Data were analysed using STATA software (version 15; StataCorp-LP). Baseline study variables were correlated with dropout status throughout the study. Variables that correlated significantly ($p < 0.05$) with dropout were entered with demographic variables into a logistic regression equation with dropout status as the independent variable.

Due to the non-normal distribution of much of the data non-parametric tests were chosen to analyse changes over time. Spearman correlations were calculated of Overall dGPCS over time with Illness Perception and PHQ variables (tables S3-5). With the exception of the PHQ scales which include overlapping items and were entered into separate equations, all correlations were under 0.75 indicating no collinearity between measures. Illness perceptions which correlated with Overall dGPCS over time ($p < 0.05$) were selected as independent variables, along with depression. Correlations showed consistency over time, with Timeline, Consequences and Emotional Representations significantly correlating with the overall outcome measure at 0, 12 and 24 months. Therefore Timeline, Consequences and Emotional Representations were selected along with depression to be entered in the logistic regression equations.

A series of logistic regressions were then undertaken. Scores on the PHQ-2 (score ≥ 3) and PHQ-4 (score ≥ 6) were used to create the dichotomous variables 'depressed' and 'psychologically distressed'. These were entered as variables into separate logistic regression equations to test their relative ability to predict outcome status (using the GPCS measure described above), controlled by demographic factors (age, gender, deprivation index (Government 2015), time since onset, diagnosis category, education level and work status). Deprivation index was derived from English government rankings of relative deprivation by neighbourhood (low ranks indicate higher deprivation).

Multiple logistic regression analyses were then carried out of the three Illness Perception subscales that correlated significantly with overall outcome, depression and the 7 demographic variables against the dependent variable overall GPCS outcome over time.

Results

Demographics of the study sample are shown in Table 1. The sample represented participants across the adult age range and from a range of social backgrounds. Females

outnumbered males by a ratio of approximately 4:1. Figure 1 demonstrates recruitment and attrition at the key time points of this study.

Just over one-third (34.33%) of participants dropped out of the study over 2 years. Dropout was not significantly associated with any of the baseline study variables, GCPS outcome category or with social class, gender, duration of pain, diagnosis, employment status or education level. However, younger participants were more likely to drop out of the study ($Z=2.11$, $p=0.035$, $OR=2.36$ 95%CI=1.06-5.23). Participants with lower initial pain scores measured by the characteristic pain intensity subscale (CPI) of the GCPS were also marginally more likely to drop out before the end of the study ($Z=-2.09$, $p=0.036$, $OR=0.981$ 95%CI=0.96-0.99). Despite patients who reported initially lower levels of pain being more likely to drop out, results of successive Wilcoxon Signed Rank tests indicated that CPI scores reduced from baseline at 12 and 24 months with significant improvements at 12 months from baseline ($Z=6.512$, $p<.000$ and at 24 months from 12 months ($Z=2.598$, $p<.010$ and from baseline ($Z=6.745$, $p<.000$).

Means and standard deviations of illness perception and mood scores at baseline are shown in table 2. Wilcoxon signed-ranks tests indicated that consequences and emotional representations differed significantly between high and low GCPS disability groups at all timepoints and timeline representations differed at 12 and 24 months (table S7). Patients who endorsed more consequences, and higher emotional impact were more likely to be in the high disability group as defined by dGCPS at all points in time. At 12 and 24 months patients with high disability also viewed pain as more likely to be permanent and as difficult to comprehend. At 24 months patients in the high disability group also reported lower levels of personal control. Other illness perceptions did not differ significantly by group.

Results of the single-factor logistic regressions of PHQ-2 and PHQ-4 against GCPS outcome are shown in Table 3. Both variables predicted significantly higher odds of a poor outcome at baseline (depression OR 4.06, psychological distress OR 3.81), 12 months (depression OR 2.94, psychological distress OR 3.81) and 24 months later (depression OR 4.91, psychological distress OR 2.13), as well as on overall GCPS outcome over time (depression OR 3.40, psychological distress OR 4.18). However, neither depression nor psychological distress remained significant when added to multifactorial models along with Illness perception variables (Table 3).

Results of controlled multiple logistic regressions are shown in table 4. Consequences ($z=3.78$, $p<0.000$, $OR 3.05$ 95%CI 1.71 – 5.43) IMD dichotomised social deprivation classification representing higher SES ($z=-2.06$, $p=0.039$, $OR 0.42$ 95%CI 0.18-0.96) and age ($z=-2.64$, $p=0.01$, $OR 0.95$ 95%CI 0.91 – 0.99) were significant predictors of overall GCPS outcome in a model that accounted for 25% of outcome variance.

At 12 months only Consequences ($z=3.31$, $p=0.001$, $OR 3.38$, 95%CI 1.64 – 6.93) and at 24 months Consequences ($z=2.70$, $p=0.007$ OR 4.09 95%CI 1.47–11.35) and age ($z=-2.19$, $p=0.028$ OR 0.93 95%CI 0.87–0.99) were the only significant variables related to poor outcome.

Discussion

The primary purpose of this study was to examine the contribution of Illness Perceptions to a GCPS overall outcome measure representing high or low disability over a period of 2 years in a cohort of patients receiving routine care for POF. We chose not to limit our definition of 'routine care' as the intention of the study was to map current practice. Our sampling strategy deliberately included patients managed both within primary and secondary care; in practice many patients moved between primary and secondary care during the course of the

study making separate analysis unrealistic. While the heterogeneity of treatment may have had an impact on study variables it also allowed us to sample a broad range of patients.

Of the Illness Representations reported, only the 'Consequences' scale of the IPQ-PS was consistently linked to our measure of outcome independent of other measures. This is in keeping with previous research (Galli et al. 2010) which reported that baseline 'Consequences' was the main significant predictor of outcome at 3 months in a similar population receiving specialist care.

These findings suggest that 'Consequences' beliefs could be the target of psychosocial interventions. Galli and colleagues (Galli et al. 2010) found that by 6 months Consequences beliefs were no longer predictive of outcome. It would be interesting to explore whether or not this change could have been related to the targeted multidisciplinary treatment reported in this study. In our sample negative perceptions of the consequences of the condition remained significantly more prevalent in people with poor outcome at 12 and 24 months and overall. It is highly unlikely that these patients accessed multidisciplinary treatment as specialist psychology and physiotherapy input for POFP was not available locally at the time of the study.

Younger age was also associated with poor outcome at 2 of the 3 data collection points, baseline and 24 months. Due to limitations discussed below it is important to be cautious about this result, however the finding does indicate that age may be relevant and should be explored further. If younger age is indeed linked with poorer outcomes this could be related to pain being more disruptive of the demands typically placed on younger people. Alternatively there could be risk factors for the development and persistence of pain which are stronger in the younger patients thus leading to the manifestation of pain at an earlier age. Given methodological limitations the one finding, at baseline that the index of multiple deprivation was also linked to overall outcome may be a chance occurrence.

In keeping with other studies (Liao et al. 2011; Velly et al. 2011; Nevalainen et al. 2017) depression was a risk factor for poor outcome in POFP. However, this effect no longer held in a controlled model including other study variables. Similar results have been reported in previous research (Velly et al. 2011) in which the significant association between depression and treatment outcome disappeared when catastrophizing was added to the model; in the current study adding a measure of consequences beliefs had the same effect. While depression is important in POFP and other chronic conditions this result underlines the importance of assessing and working with depression in context. A thorough formulation of factors which are instrumental in maintaining depression, leading to a specific treatment plan, is important as standardised generic treatments for depression may miss important individual differences or factors specific to POFP or more broadly to chronic conditions.

Given that perceptions of 'Consequences' do seem to be important it is worth looking further at what this scale measures. The version of the scale used in this study includes three questions:

- My pain has major consequences on my life
- My pain causes difficulties for those who are close to me
- My pain strongly affects the way others see me.

These questions largely refer to beliefs about the impact of pain in the social domain. This domain of the biopsychosocial model is the least well researched in relation to pain management, although there is evidence that physical and social pain may share the same pathways (Eisenberger and Lieberman 2004). Social factors precipitated by pain may

include changes in relationships caused by factors such as misunderstanding, shame and stigma (Karos et al. 2018). This is consistent with descriptions given by patients of the impact of their POFP (Durham et al. 2010). The 'Consequences' scale used in this study may be capturing participants' understanding of their experience in this domain.

It seems important that a biopsychosocial assessment of POFP should include an assessment of Consequences beliefs. The relevance of these beliefs in our study seems to support the assertion (Karos et al. 2018) that routine assessment should include 'standard assessment of interpersonal needs and motivations' and that treatment should aim to recognise and address these as far as possible. In terms of psychosocially focused treatments such as cognitive behavioural therapy (CBT) and self-management a comprehensive assessment and formulation may help to determine which treatment approaches are most likely to be of help. While traditional CBT interventions can be good for targeting erroneous beliefs, thoughts about the social impact of pain are likely to include an element of truth and it would be difficult to assess their accuracy. Third-wave CBT interventions such as Acceptance and Commitment Therapy (Hayes and Wilson 1994; McCracken et al. 2004), or Compassion Focused Therapy (Gilbert 2009; Penlington 2018), may be better placed to target these beliefs and related behaviours.

The secondary aim of this study was to test the relevance of the PHQ-4 and the PHQ-2 in this population. Scores above the recommended clinical cut off on both measures (Löwe et al. 2010) were associated with significantly increased odds of poor outcome. These odds were no longer significant when Illness perceptions were added into the models, indicating that there may be some shared variance between the constructs measured by the instruments (PHQ-2 and-4 and IPQ-PS).

Depression has been identified as a risk factor for onset or poor outcome in POFP in a number of previous prospective studies using well-validated more extensive instruments resulting in odds ratios varying between 1.67 and 2.5 (Liao et al. 2011; Velly et al. 2011; Nevalainen et al. 2017). The odds ratios in the current study of between 2.94 and 4.06 related to the PHQ-2 and PHQ-4 dichotomised measures seem to be largely consistent with these studies. However, these findings are limited by wide confidence intervals and the lack of an equivalent but more extensive measure against which they could be validated. While further research is needed, this study supports the use of the PHQ-4 as a useful screening measure in POFP with a cut off score of 6, in line with similar findings in other populations (Kroenke et al. 2009; Löwe et al. 2010).

There are some limitations to this study which include the dropout rate of 34% over the two years of the study. This may limit the findings, particularly with regard to younger patients and those who had lower levels of pain who were more likely to drop out. In terms of other variables the lack of significant differences at baseline between those who stayed in the study or dropped out suggest that any bias caused by study dropout may be limited.

The drop-out rate may also affect the logistic regression where the limitation of sample size is the smaller of the two outcome groups (Babyak 2004). In this study this was $n=77$ at baseline, reducing to $n=35$ at 24 months, suggesting that results might be artificially inflated in equations including more than 3 variables. This said the selection of Consequences Beliefs was a priori based on previous research and this and the consistency of its relationship with a poor outcome over time (both single and multiple factor models) strengthens the possibility that this is a true finding. Other findings reported which occur less consistently must be treated with a greater degree of caution and be subject to further research which takes into account the likelihood of a high degree of attrition and an uneven allocation between outcome categories.

The allocation of patients, over a period of two years and five separate data collection points, to a single outcome category required some compromises to be made. Overall outcome was defined as the mode outcome over the 2 year period, ie 3 or more data points at which a 'high' or 'low' disability GCPS score was collected. The rationale for this was that POFP is a long term condition which naturally fluctuates and that a composite measure over time was therefore most likely to be a more accurate representation of long term outcome. This method cannot, however, detect changes over time that may be associated with treatment or other factors. For the variables considered in this study the consistency of findings reported using the dGCPS measure at baseline, 12 and 24 months and those using the overall outcome measure suggest that, despite the limitations of the overall outcome measure the findings remain robust.

The lack of inclusion of a questionnaire which has been validated in this population specifically for depression against which to compare the PHQ-4 weakened our ability to thoroughly test the utility of this measure. It would be beneficial for future research to include a well validated depression questionnaire such as the Patient Health Questionnaire-9 (PHQ-9) for a more robust exploration of the performance of the PHQ-4.

Strengths of the study, however, include the longitudinal timeframe of 2 years which allowed for data to be collected on at least 3 (and for some measures 5) separate occasions. The main results reported remained consistent across more than one point in time which adds confidence to the findings. The broad recruitment strategy which included a range of patients with POFP across both primary and secondary care settings also adds strength to the generalizability of our findings.

Conclusion

This study has demonstrated that patients' perceptions of the Consequences of their POFP are important in a community sample of people receiving routine care but that other Illness Perceptions may be of little predictive value. It has also provided some initial support for the use of the PHQ-4, an easy to administer questionnaire consisting of just four questions to screen for depression and anxiety in this population.

While it has previously been suggested that the GCPS can be used to target the appropriate level of care for patients (Durham et al. 2016), the addition of a brief assessment of key biopsychosocial measures may also help with targeted treatment. This could be used to triage patients who may benefit from a supported psychological or self-management intervention alongside their dental care. In the related area of back pain a brief questionnaire consisting of only 9 items has been successful in targeting treatment, reducing costs and improving outcomes (Hill et al. 2011). Further work would be helpful to explore whether a similar approach could be successful for orofacial pain management.

Acknowledgements

This study and J.D. are funded by the National Institute for Health Research (NIHR; Clinician Scientist Award NIHR-CS-011-003). We thank the patients, health professionals and steering group members for their valuable time and assistance. The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

References

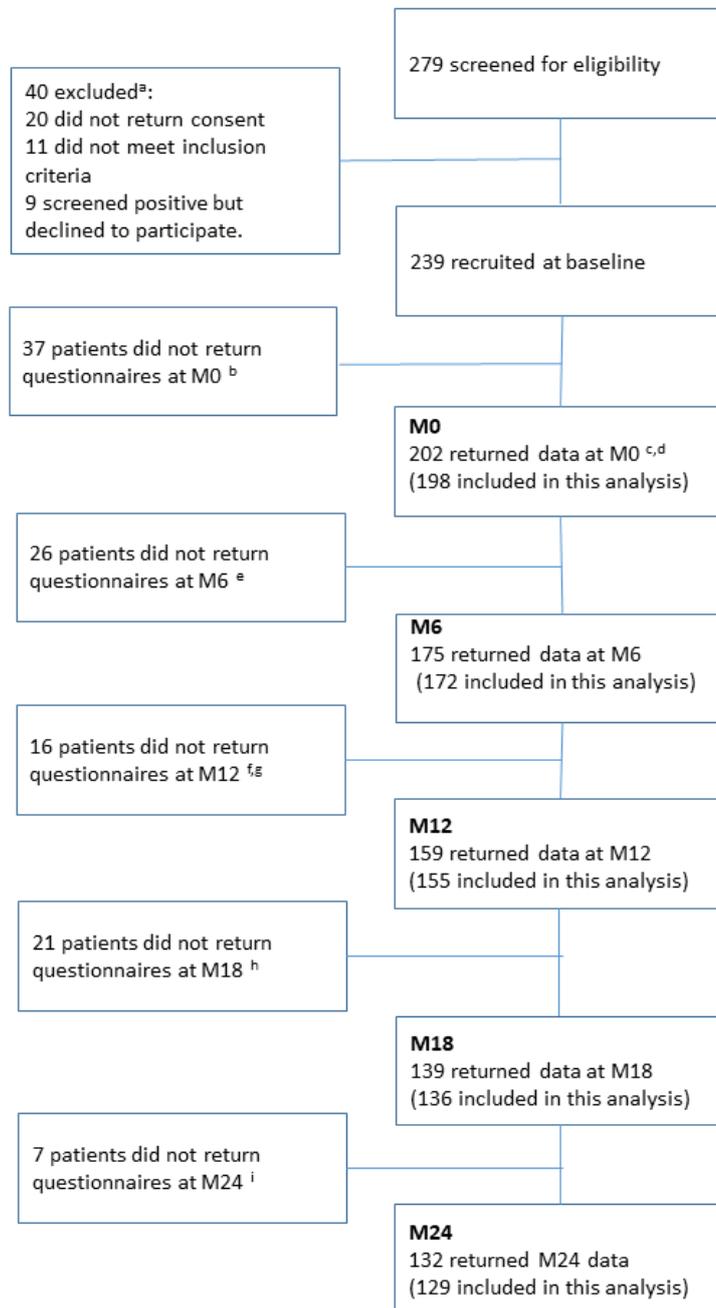
- Aggarwal VR, Macfarlane GJ, Farragher TM, McBeth J. 2010. Risk factors for onset of chronic orofacial pain—results of the North Cheshire orofacial pain prospective population study. *PAIN*® 149(2):354-359.
- APA. 2013. Diagnostic and statistical manual of mental disorders (DSM-5®) American Psychiatric Pub.
- Babak MA. 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine* 66(3):411-421.
- Beecroft EV, Durham J, Thomson P. 2013. Retrospective examination of the healthcare 'journey' of chronic orofacial pain patients referred to oral and maxillofacial surgery. *Br Dent J* 214(5):E12.
- Bonathan CJ, Zakrzewska JM, Love J, de C W, Amanda C. 2014. Beliefs and distress about orofacial pain: patient journey through a specialist pain consultation. *Journal of Oral & Facial Pain & Headache* 28(3).
- Breckons M, Shen J, Bunga J, Vale L, Durham J. 2018. DEEP Study: Indirect and Out-of-pocket Costs of Persistent Orofacial Pain. *Journal of dental research*:0022034518773310.
- Cohen J. 1992. A power primer. *Psychological bulletin* 112(1):155.
- Durham J, Steele J, Wassell R, Exley C. 2010. Living with uncertainty: temporomandibular disorders. *Journal of dental research* 89(8):827-830.
- Durham J, Breckons M, Araujo-Soares V, Exley C, Steele J, Vale L. 2014. Developing Effective and Efficient care pathways in chronic Pain: DEEP study protocol. *BMC oral health* 14(1):6.
- Durham J, Shen J, Breckons M, Steele JG, Araujo-Soares V, Exley C, Vale L. 2016. Healthcare Cost and Impact of Persistent Orofacial Pain: The DEEP Study Cohort. *J Dent Res* 95(10):1147-1154.
- Dworkin SF. 1994. Perspectives on the interaction of biological, psychological and social factors in TMD. *Journal of the American Dental Association (1939)* 125(7):856-863.
- Dworkin SF, Huggins KH, Wilson L, Mancl L, Turner J, Massoth D, LeResche L, Truelove E. 2002. 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* 16(1).
- Eisenberger NI, Lieberman MD. 2004. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in cognitive sciences* 8(7):294-300.
- Galli U, Ettlin DA, Palla S, Ehlert U, Gaab J. 2010. Do illness perceptions predict pain-related disability and mood in chronic orofacial pain patients? A 6-month follow-up study. *Eur J Pain* 14(5):550-558.
- Gilbert P. 2009. Introducing compassion-focused therapy. *Advances in psychiatric treatment* 15(3):199-208.
- Government E. 2015. The English Index of Multiple Deprivation. In: Government DfCaL editor.
- Hayes SC, Wilson KG. 1994. Acceptance and commitment therapy: Altering the verbal support for experiential avoidance. *The Behavior Analyst* 17(2):289-303.
- Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, Konstantinou K, Main CJ, Mason E, Somerville S. 2011. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *The Lancet* 378(9802):1560-1571.
- Karos K, Williams ACdC, Meulders A, Vlaeyen JW. 2018. Pain as a threat to the social self: a motivational account. *Pain*.
- Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. 2007. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of internal medicine* 146(5):317-325.
- Kroenke K, Spitzer RL, Williams JBW, Löwe B. 2009. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics* 50(6):613-621.

- Liao CH, Chang CS, Chang SN, Lane HY, Lyu SY, Morisky DE, Sung FC. 2011. The risk of temporomandibular disorder in patients with depression: a population-based cohort study. *Community Dent Oral Epidemiol* 39(6):525-531.
- Litt MD, Shafer DM, Kreutzer DL. 2010. Brief cognitive-behavioral treatment for TMD pain: long-term outcomes and moderators of treatment. *Pain* 151(1):110-116.
- Litt MD, Porto FB. 2013. Determinants of pain treatment response and nonresponse: identification of TMD patient subgroups. *J Pain* 14(11):1502-1513.
- Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, Schneider A, Brähler E. 2010. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of affective disorders* 122(1):86-95.
- McCracken LM, Carson JW, Eccleston C, Keefe FJ. 2004. Acceptance and change in the context of chronic pain. *Pain* 109(1):4-7.
- Moss-Morris R, Weinman J, Petrie K, Horne R, Cameron L, Buick D. 2002. The revised illness perception questionnaire (IPQ-R). *Psychology and health* 17(1):1-16.
- National Collaborating Centre for Mental Health. 2018. The Improving Access to Psychological Therapies (IAPT) Pathway for People with Long-term Physical Health Conditions and Medically Unexplained Symptoms. Full Implementation Guidance. In: Health NCCfM editor. London.
- Nevalainen N, Lähdesmäki R, Mäki P, Ek E, Taanila A, Pesonen P, Sipilä K. 2017. Association of stress and depression with chronic facial pain: A case-control study based on the Northern Finland 1966 Birth Cohort. *CRANIO®* 35(3):187-191.
- Penlington C. 2018. Exploring a compassion-focused intervention for persistent pain in a group setting. *British Journal of Pain*:2049463718772148.
- Pincus T, Burton AK, Vogel S, Field AP. 2002. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 27(5):E109-E120.
- Shueb S, Nixdorf D, John M, Alonso BF, Durham J. 2015. What is the impact of acute and chronic orofacial pain on quality of life? *Journal of dentistry* 43(10):1203-1210.
- Sniehotta FF, Gorski C, Araújo-Soares V. 2010. Adoption of community-based cardiac rehabilitation programs and physical activity following phase III cardiac rehabilitation in Scotland: A prospective and predictive study. *Psychology and Health* 25(7):839-854.
- Velly AM, Look JO, Carlson C, Lenton PA, Kang W, Holcroft CA, Friction JR. 2011. The effect of catastrophizing and depression on chronic pain--a prospective cohort study of temporomandibular muscle and joint pain disorders. *Pain* 152(10):2377-2383.
- Von Korff M, Ormel J, Keefe FJ, Dworkin SF. 1992. Grading the severity of chronic pain. *Pain* 50(2):133-149.

Number of references - 34

Figure 1. Consort Diagram showing recruitment and attrition at key study time points.

The diagram below shows overall dropout from the study at each time point and also indicates the number of people included in this analysis.



- There was no significant difference in age, gender or origin of pain between positively screened patients who participated and those who declined ($P>0.05$).
- There were no significant differences in gender, ethnicity, duration of pain or origin of pain between those dropping out and the 198 participants whose data was included in the study ($p>0.05$). However those participating were significantly older than those who dropped out ($p<0.01$).
- One patient withdrew from the study at M18 and requested that their data be withdrawn.
- Data was not received from 3 participants at M0 who reported returning data but this was not received by the study team. These patients returned data at subsequent time points.
- There was no significant difference between those dropping out at M6 and the M0 sample on the basis of: Age, Gender, Ethnicity, Duration of pain or Origin of pain ($P>0.05$).
- There was no significant difference between those dropping out at M12 and the M0 sample on the basis of: Age, Gender, Ethnicity or Origin of pain ($P>0.05$) although those dropping out had a significantly longer duration of pain to those participating at baseline ($P<0.05$).
- One of these patients went on to complete data at a further time point.
- There was no significant difference between those dropping out at M18 and the M0 sample on the basis of: Age, Gender, Duration of pain or Origin of pain ($P>0.05$), although those dropping out consisted of a greater proportion of a White British ethnic group than those participating at baseline ($P<0.01$).
- There was no significant difference between those dropping out at M24 and the M0 sample on the basis of: Age, Gender, Duration of pain or Origin of pain ($P>0.05$), although those dropping out consisted of a greater proportion of a White British ethnic group than those participating at baseline ($P<0.01$).

Table 1: Demographic variables at baseline

Variable n	Male	Female	Total
Years in pain:			
<1 year	2	25	27
1 – 5 years	12	70	82
>5 years	24	65	89
Origin of pain			
Musculoskeletal	18	68	86
Neuropathic/vascular	11	53	64
Combined	9	39	48
Deprivation index*			
Bottom 50%	21	60	81
Top 50%	17	97	114
Employment			
Employed	12	77	89
Retired	18	52	70
Other	7	29	36
Age			
20 – 39	5	41	46
40 – 59	13	64	77
60+	20	52	72
GCPS grade			
0	0	3	3
1	12	49	61
2a	12	45	57
2b	5	30	35
3	5	26	31
4	4	7	11
GCPS category			
Low	24	97	121
High	14	63	77

Breakdown of demographic information by gender for full study sample (n=198).

*Deprivation index refers to classification of home neighbourhood by the English Index of Multiple Deprivation (IMD) (1) into deciles of deprivation which are then further collapsed into categories of high deprivation and low deprivation.

Table 2: Mean (SD) self-report scores on illness perceptions and mood at baseline by gender

	Male		Female		Overall	
Timeline	3.79	(0.92)	3.61	(0.85)	3.64	(0.86)
Consequences	3.18	(1.22)	2.91	(1.03)	2.96	(1.07)
Personal Control	2.80	(0.90)	3.06	(0.79)	3.01	(0.82)
Treatment Control	2.99	(0.80)	3.08	(0.85)	3.06	(0.84)
Cyclical Timeline	3.78	(0.74)	3.49	(0.96)	3.55	(0.93)
Illness Coherence	3.09	(1.06)	3.27	(1.16)	3.23	(1.14)
Emotional Representations	3.28	(1.15)	3.20	(1.10)	3.22	(1.10)
PHQ4	3.97	(3.51)	3.36	(3.48)	3.48	(3.48)
PHQ2	2.16	(2.12)	1.5	(1.81)	1.63	(1.89)
GAD2	1.82	(1.71)	1.84	(1.99)	1.83	(1.94)

Illness perceptions include 7 subscales of the IPQ-PR as shown. PHQ2 (depression) and GAD2 (anxiety) scales combine to form the PHQ4. Means and standard deviations reported for the full sample (n=198) and separately for males (n=38) and females (n=160). Mann-Whitney tests indicated no significant differences between males and females on any of the above scales.

Table 3: Logistic regression of PHQ-2 and PHQ-4 at baseline as predictors of poor outcome over time.

	PHQ-2			PHQ-4		
	OR	95% CI	p	OR	95% CI	p
Baseline	4.06	1.74 – 9.52	.001	2.88	1.12 – 7.41	.029
12 months	2.94	1.03 – 8.42	.044	3.81	1.15 – 12.58	.028
24 months	4.91	1.33 – 18.19	.017	2.13	0.52 – 8.69	.294
Overall	3.40	1.41 – 8.17	.006	4.18	1.50 – 11.67	.006

Results of single-factor logistic regressions controlled for demographic factors (age, gender, deprivation index, time since onset, diagnosis category, education level and work status)

Table 4: Collated results of three logistic regressions showing odds ratios of baseline variables[†] regressed against dGCPS over time

Independent variable at baseline (M0) [†]	Odds ratio of overall good outcome over time (dGCPS) [‡]					
	Baseline (n=167 obs.)		12 months (n=131 obs.)		24 months (n=104 obs.)	
	OR	p	OR	p	OR	p
Timeline	0.907	0.727	0.972	0.933	1.507	0.315
(95% Confidence Intervals)	0.521 – 1.578		0.504 – 1.873		0.677 – 3.354	
Consequences	3.048	0.000**	3.376	0.001**	4.086	0.007**
(95% Confidence Intervals)	1.710 – 5.431		1.643 – 6.933		1.471 – 11.348	
Emotional Representations	0.771	0.291	1.037	0.919	2.091	0.159
(95% Confidence Intervals)	0.475 – 1.249		0.517 – 2.077		0.750 – 5.829	
PHQ-2	1.384	0.533	2.063	0.271	0.572	0.547
(95% Confidence Intervals)	0.499 – 3.837		0.568 – 7.488		0.0993 – 3.519	
Index of multiple deprivation (IMD)	0.417	0.039*	0.802	0.689	0.342	0.132
(95% Confidence Intervals)	0.182 – 0.958		0.271 – 2.368		0.085 – 1.384	
Age	0.951	0.008**	0.959	0.088	0.930	0.028*
(95% Confidence Intervals)	0.917 – 0.987		0.914 – 1.006		0.871 – 0.992	

[†]Results of multiple logistic regression equations of variables which initially correlated at $p < 0.05$ with dichotomised Graded Chronic Pain Scale d(GCPS) [‡] outcome measure at 12 months, 24 months and overall.

Significance is indicated by *($p < 0.05$) or **($p < 0.01$)

[‡]dGCPS: M12 & M24 were low vs high dGCPS state; overall GCPS outcome was 'poor' if ≥ 3 high states over time and 'good' otherwise i.e. the mode state was taken across the five timepoints

Predicting Persistent Orofacial Pain: the role of illness perceptions, anxiety and depression.

Penlington, C., Araujo-Soares, V. and Durham, J.

Supplementary Information

Supplementary Methods

Recruitment took place over a period of 22 months and included prospective and retrospective recruitment from primary care (non-specialist treatment) and prospective recruitment within secondary care (specialist treatment). For retrospective recruitment specific diagnostic codes referring to persistent orofacial pain were identified and patient records in primary care electronically searched for a match to one or more of these codes during the preceding 12 months. Patients identified using this method were contacted in writing and invited to make contact with the research team if they were interested in participating in the study. Patients recruited prospectively were approached by a member of staff, given a short description of the study and asked if they were interested in participating. All patients who expressed an interest in participating were given a study pack including an information sheet, initial consent form and a self-report screening measure. They were then contacted by telephone by the research team who completed the screening measure over the telephone. Those screening positive and giving informed consent were enrolled into the study.

Once enrolled in the study following informed consent and a positive screening result structured interviews were then completed with a trained interviewer at baseline. Follow-up data was collected by post. Study questionnaires were mailed to participants at each study timepoint, and a period of 14 days allowed for their return. Patients with missing or incomplete data after this period were followed up by telephone to remind them to complete the questionnaires or clarify any missing or confusing data.

Repeated measures Friedman's two-way ANOVAs were performed on each of the 7 illness perceptions tested to monitor change over time.

Supplementary Results

A total of 387 individuals were referred for screening for eligibility for the study. Of these 12 were from retrospective recruitment (3%). The majority (72%) of those referred accepted the invitation to be screened with two individuals from retrospective recruitment declining to be screened. There was no significant difference in gender ($\chi^2(1, n=386)=0.66$; $p=0.261$), age ($t(366)=1.24$; $p=0.215$; 95%CI difference -1.52, 6.73years), recruitment ($\chi^2(1, n=386)=0.66$; $p=0.261$), between those declining and accepting the invitation to be screened. There were a significantly greater number of declined invitations from those referred from primary care than those referred from secondary care ($\chi^2(1, n=386)=6.610$; $p=0.01$). Figure 1 demonstrates recruitment and attrition at the key time points of this study.

Perceptions of consequences ($Z=3.079$, $p<.01$) and emotional representations $Z=3.365$, $p<.01$) scores changed over time, showing some improvement from baseline to 24 months (table S6). There was no significant change over time in perceptions of timeline, personal or treatment control, illness coherence or cyclical timeline.

When entered as single factors, Consequences and Emotional representations consistently predicted dGCPS grouping at every point in time (data not shown). Since Consequences beliefs were consistently related to outcome, these were further explored with Spearman correlations. Correlations at baseline were carried out between Consequences beliefs, characteristic pain index, PHQ4, age, classification of pain, index of deprivation, education and employment status. The same correlations were repeated at 12 and 24 months (using baseline measures of age, classification of pain, index of deprivation, education and employment status as data on these measures were only collected once). Results are shown in table S8 Significant correlations were reported between Consequences beliefs and younger age, PHQ4 score, classification of pain, and being neither employed nor retired. Index of deprivation and employment status were not correlated with Consequences beliefs.

The reported correlations of 'Consequences' beliefs with higher reported pain younger age, higher psychological distress, unemployment and reduced likelihood of a simple 'musculoskeletal' diagnosis of pain may be relevant in terms of understanding the impact that POFP has in different circumstances on many aspects of somebody's life.

Supplementary Table 1 (S1): Illness Perceptions Questionnaire-Psychometrically Shortened (IPQ-PS)

	Views about your pain	Relevant Domain
1	I don't understand my pain	Illness Coherence
2	My pain will last for a long time	Timeline
3	I get depressed when I think about my pain	Emotional Representations
4	My pain has major consequences on my life	Consequences
5	My treatment will be effective in curing my pain	Treatment Control
6	Having this pain makes me feel anxious	Emotional Representations
7	Nothing I do will affect my pain	Personal Control
8	My pain doesn't make any sense to me	Illness Coherence
9	The negative effects of my pain can be prevented by my treatment	Treatment Control
10	My pain is very unpredictable	Timeline Cyclical
11	My pain is a mystery to me	Illness Coherence
12	My pain is likely to be permanent rather than temporary	Timeline
13	My pain causes difficulties for those who are close to me	Consequences
14	I have the power to influence my pain	Personal Control
15	My symptoms come and go in cycles	Timeline Cyclical
16	My pain strongly affects the way others see me	Consequences
17	My treatment can control my pain	Treatment Control
18	I go through cycles in which my pain gets better and worse	Timeline Cyclical
19	When I think about my pain I get upset	Emotional Representation
20	I expect to have this pain for the rest of my life	Timeline
21	My actions will have no effect on the outcome of my pain	Personal Control

IPQ-PS scale which was used in the study to assess Illness Representations. Items are scored on a scale of 1 (strongly disagree) to 5 (strongly agree).

Supplementary Table 2 (S2): Missing data for each measure at baseline, 12 and 24 months.

Measure	Data collection timepoint		
	Baseline (n=198)	12 months (n=155)	24 months (n=129)
GCPS* (%)	3 (1.49%)	3 (1.91%)	0 (0%)
PHQ4** (%)	11 (5.47%)	7 (4.46%)	13 (9.42%)
IPQ-PS** (%)	6 (2.99%)	8 (5.10%)	9 (6.52%)

Table shows number of participants in study with missing data for each measure at each time of data collection. *GCPS data was also collected at 6 months (n=172, 1 missing (0.6%)) and 18 months (n=136, 2 missing (1.4%)). **PHQ4 and IPQ-PS data was collected at baseline, 12 and 24 months only.

Supplementary Table 3 (S3): Correlations at Baseline between all measures.

Table shows Spearman correlations at baseline between Overall dichotomised Graded Chronic Pain Scale over time outcome category (dGCPS) and Illness perceptions subscales and PHQ subscales at baseline. PHQ-4 includes items from PHQ-2 (these scales were never

	Time	1	2	3	4	5	6	7	8	9	10
Graded Chronic Pain Scale category (1)		0.209**	0.333**	-0.012	-0.123	-0.015	-0.032	0.181*	0.309**	0.235**	
timeline (2)			0.457**	-0.206**	-0.327**	-0.097	-0.059	0.455	0.296**	0.272**	
consequences (3)				-0.165*	-0.163*	-0.204**	0.055	0.692**	0.431**	0.430**	
personal control (4)					0.3417**	0.404**	0.001	-0.342**	-0.223**	-0.160*	
illness coherence (5)						0.140	0.160*	-0.159*	-0.071	-0.044	
treatment control (6)							-0.48	-0.47	-0.198**	-0.168*	
cyclical timeline (7)								0.007	-0.027	0.085	
emotional representations (8)									0.518**	0.518**	
PHQ-2 (9)										0.883**	
PHQ-4 (10)											
p		0.004	0.000	0.872	0.835	0.090	0.664	0.012	0.000	0.000	

included together in the same equation). Significance: * indicates $p < .05$, ** $p < .01$. Bottom row shows p values at baseline of correlations of each measure with the dGCPS dichotomised outcome measure.

Supplementary Table 6 (S6): Changes over time in illness perceptions and characteristic pain index measured by Wilcoxon signed-rank test.

Illness perception	Difference (Z)	Baseline mean (SD)	24month mean (SD)
Timeline	0.988	3.8 (±0.9)	3.7 (±1)
Consequences	3.079**	3.0 (±1.1)	2.8 (±1.2)
Personal Control	-1.530	3.1 (±0.8)	3.2 (±0.9)
Treatment Control	0.237	3.1 (±0.9)	3.0 (±1.0)
Illness Coherence	-2.043	3.3 (±1.2)	3.5 (±1.2)
Cyclical Timeline	0.238	3.5 (±1.0)	3.6 (±0.9)
Emotional Representations	3.365**	3.3 (±1.1)	3.1 (±0.9)
Characteristic pain intensity	6.745**	55.3 (±22.0)	40.7 (±25.0)

** p<.01

Supplementary Table 7 (S7): Differences in illness perceptions by outcome category across time.

	Baseline				
Illness perceptions construct mean score (SD):	Low GCPS (n=121)	High GCPS (n=77)	Total (n=198)	Difference (z)	p
Timeline	3.7(±0.8)	3.9(±1)	3.8(±0.9)	-1.778	0.075
Consequences	2.6(±1)	3.6(±1)	3(±1.1)	-6.469	0.000**
Personal Control	3(±0.9)	3.1(±0.8)	3.1(±0.8)	-0.879	0.379
Treatment control	3.3(±0.9)	2.9(±0.8)	3.1(±0.9)	1.404	0.160
Illness coherence	3.3(±1.2)	3.1(±1.2)	3.3(±1.2)	0.624	0.533
Timeline cyclical	3.5(±1)	3.6(±1)	3.5(±1)	-0.542	0.588
Emotional representations	3(±1.2)	3.8(±0.8)	3.3(±1.1)	-4.300	0.000**
	12 months				
	Low GCPS (n=113)	High GCPS (n=43)	Total (n=156)	Difference (z)	p
Timeline	3.6(±0.9)	4.2(±0.7)	3.8(±0.9)	-3.772	0.000**
Consequences	2.5(±1)	4(±0.9)	3(±1.2)	-6.815	0.000**
Personal Control	3.2(±0.8)	3(±0.8)	3.2(±0.8)	1.541	0.123
Treatment control	3.1(±0.8)	2.6(±0.7)	3(±0.8)	2.823	0.005**
Illness coherence	3.6(±1.1)	2.9(±1.2)	3.4(±1.2)	3.608	0.000**
Timeline cyclical	3.5(±0.9)	3.5(±0.9)	3.5(±0.9)	0.049	0.961
Emotional representations	3(±1)	4(±0.9)	3.3(±1.1)	-5.437	0.000**
	24 months				
	Low GCPS (n=96)	High GCPS (n=35)	Total (n=131)	Difference (z)	p
Timeline	3.4(±1)	4.3(±0.7)	3.7(±1)	-4.699	0.000**
Consequences	2.3(±0.9)	3.8(±0.9)	2.8(±1.2)	-6.398	0.000**
Personal Control	3.3(±0.9)	2.9(±0.8)	3.2(±0.9)	2.582	0.010**
Treatment control	3.1(±1)	2.9(±1)	3(±1)	1.576	0.115
Illness coherence	3.6(±1.1)	3.1(±1.2)	3.5(±1.2)	3.608	0.000**
Timeline cyclical	3.6(±0.9)	3.4(±0.8)	3.6(±0.9)	1.352	0.176
Emotional representations	2.7(±1)	3.8(±0.9)	3.1(±1.1)	-5.018	0.000**

Table shows means (standard deviations) of Illness Perception scales at baseline, 12 months and 24 months by dichotomised Graded Chronic Pain Scale (dGCPS) (Dworkin et al. 2002) indicating high or low disability at each point in time and significance of differences

between high and low disability groups. Wilcoxon signed-rank tests were used to test for difference between low and high GCPS outcome status for each variable and at each point in time at $**p<.01$.

Supplementary Table 8 (S8): Spearman correlations between Consequences, characteristic pain intensity, PHQ4 and demographic factors

	Baseline	12 months	24 months
Age	-0.182*	-0.241**	-0.309**
Characteristic pain intensity	0.331**	0.436**	0.541**
PHQ4	0.394**	0.405**	0.387**
Male	0.062	0.021	0.015
Musculoskeletal diagnosis	-0.261**	-0.285**	-0.344**
Neuropathic diagnosis	0.216**	0.159	0.218**
Combined diagnosis	0.069	0.177*	0.177
Employed	-0.024	0.028	0.120
Retired	-0.137	-0.101	-0.185
Not employed or retired	0.198*	0.095	0.086

Table shows Spearman correlations between Consequences scores, characteristic pain intensity and PHQ4 scores reported at each time of data collection. Demographic factors were measured only at baseline so correlations including these factors used the baseline measures. No significant correlations were reported between Consequences and Deprivation Index or level of education (data not shown).

Significant correlations are indicated by *(p<.05) and ** (p<.01).