School of Physiotherapy

Topical menthol identifies cold hyperalgesia in individuals with chronic pain from knee osteoarthritis

Penny Moss

This thesis is presented for the Degree of Doctor of Philosophy of Curtin University

March 2013
Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Penny Moss
March 2013
Acknowledgments

Every PhD is a team effort and mine is no exception. There are so many different people who have assisted in so many different ways in order to make its completion possible, but I would like to try to acknowledge at least some of them.

Above all there are the study participants, without whose generosity of time and effort there would be no numbers in spread-sheets to analyse and no conclusions to be drawn. The majority of subjects participate with no expectation of any personal benefit other than a sense that they have assisted the scientific process, so that there is a real responsibility to ensure that these results are published and acted upon.

There is inevitably funding involved in the completion of a PhD study. For this study I would like to acknowledge the assistance of a research grant from Merck, Sharp and Dohme, which funded my time for completion of the osteoarthritis studies in this investigation. In addition, I could not have completed the final thesis element in any reasonable time-frame without the assistance of six months PhD completion leave provided by Curtin University. So thank you.

Associated with this, my work colleagues must be given an immense thank you for the support provided over the several years of this investigation and final thesis completion. Not only have colleagues uncomplainingly taken over my work responsibilities whilst I have been on leave but they have respected the seclusion that is needed to write effectively. Other colleagues have offered a supportive chat or email from time to time, all of which I have very much appreciated during a process that can be very isolating.

Outside work, there are inevitably a wide range of people who have had to modify their expectations for the sake of PhD completion. Thank you to my family in UK for managing some very difficult issues this past year without expecting me to drop everything, but instead providing me with constant support and encouragement. For my extended family in Australia, thank you for the constant, unwavering support and understanding, from the 3-year old accepting that I often couldn’t play because I was writing my ‘book’, to the formatting consultants, sounding boards and test guinea pigs!
Enormous thanks of course to my supervisors Tony and Heather for being there to bounce ideas off, for keeping me on track but allowing me to work within my own time-frame and above all for managing the complex task of balancing the roles of supervisors, work colleagues, research colleagues and friends. This is truly a collaborative work.

Then to my four-legged supervisors, Jasper and Charlie, who have patiently accompanied me through the final writing process, sitting by my feet, listening to my explanations of complex neurophysiology and not arguing with me, sending me calming ‘vibes’ and only hassling me when it was time for a walk or for food.

Finally, the biggest thanks of all goes to Chris, without whom this thesis would still be in the making, the house would be falling apart, the cupboard would be bare and I would be malnourished. Even though I have made your life considerably more stressful, you have unwaveringly been my sounding board, my advisor, my organizer, and my constant encourager. I could not have completed this without you. Thank you, thank you ....
Abstract

Background and Aims
Chronic pain is a world-wide problem, affecting one in five adults and associated with high societal and personal costs. A chronic pain diagnosis is currently made by longevity of unresolved symptoms and failure to respond to standard interventions. However, by the time chronic pain has been identified, pain processes will have become entrenched and secondary physical, psychological and social consequences developed. Pain resolution becomes increasingly difficult. A test that enabled earlier identification of individuals who are at risk of developing chronic pain would allow more appropriately targeted interventions to be provided before secondary effects were established. Increasing evidence suggests that the development of chronic pain is associated with centrally-driven pain augmentation rather than specific patho-etiologies. Widespread mechanical and cold hyperalgesia, reduced spinal inhibition and decreased thalamic volume have all been reported for individuals with the highest pain and dysfunction from both fibromyalgia and painful knee osteoarthritis (OA). For these individuals pain is driven by central rather than peripheral mechanisms.

Presence of cold hyperalgesia may be a useful clinical indicator of centrally-augmented pain. Cold hyperalgesia is a key indicator of severity in widespread pain syndromes and is strongly associated with thalamic changes. Elevated cold pain threshold and changes to the quality of cold response have also been associated with dysfunction in supraspinal inhibitory processes. However conventional assessment of cold hyperalgesia using a peltier thermode is problematic from both an equipment and reliability perspective. Topical menthol is an established chemical agent that activates TRPM8, a receptor channel linked to cold transduction. The current investigation therefore hypothesised that topical menthol might provide a more reliable and valid alternative test for cold hyperalgesia. The purpose of this investigation therefore was to develop a new, clinically-applicable topical test for cold hyperalgesia and to evaluate the sensitivity of this test in identifying individuals with additional signs of centrally-augmented pain associated with painful knee OA.

Method
A series of seven studies evaluated the response characteristics, validity and reliability of a menthol test developed during preliminary laboratory investigations (Appendices 1 and 2). All subjects were volunteers from the Perth community, recruited by general advertising and word of mouth. The same basic application method was used for each of the studies: the same volar forearm test site was used and the menthol formulation
always applied to cleaned skin and occluded by a Tegaderm adhesive dressing. Sensory response was always quantified by an individual rating of sensation intensity and quality at regular time-points during the 15-minute application.

Studies 1 and 2 evaluated validity and reliability of graded menthol formulations in 32 and 27 healthy individuals, with no history of chronic pain. Studies 1 (Chapter 3) and 2 (Chapter 4) used a blinded, randomised, repeated measures design to assess the dose-dependent effects of graded concentrations of menthol, initially in solution (Chapter 3) and then in a gel formulation (Chapter 4), which provided more consistent topical delivery. Study 1 evaluated concurrent criterion validity by comparing menthol response to peltier-thermode tested CPT values at the same forearm site. Study 2 also characterised normal and atypical responses and determined which gel concentration showed the greatest discriminatory ability. Response to menthol during the 15-minute application was measured by: 1) VAS intensity of cold, heat, unpleasantness and pain sensation every one-minute; 2) selection of descriptor(s) from a McGill pain descriptors list every two minutes, quantified as McGill Pain Rating Index (PRI). Study 3 evaluated in 29 healthy subjects the validity and sensitivity of a newly developed measurement index to identify a cold hyperalgesic response (the Algotect Descriptor Index (ADI)), combining weighted VAS and descriptor data into a single value. The index was applied to measure response to three sustained graded cold-temperature stimuli, each applied to the volar forearm site for 5 minutes using a peltier thermode. ADI scores were compared for temperature-dependent effects alongside CPT values for additional criterion validity. Study 4 used a test-retest design to evaluate the reliability of the menthol test in healthy subjects over two test sessions.

Chapters 7 to 9 applied the menthol test to subjects with chronic pain from knee OA, evaluating whether it was able to identify individuals who exhibited additional signs of centrally-augmented pain. OA volunteers were assessed by a Rheumatologist as having radiographic evidence of knee OA, knee pain >3/10, and no history of other chronic pain. Studies 5 and 6 evaluated the sensitivity of the menthol ADI test in identifying those with cold hyperalgesia when assessed alongside conventional CPT. These studies also investigated the association between high ADI and elevated response to other measures related to centrally-augmented pain: quantitative sensory tests at local and distant sites and self-report measures of presence of neuropathic pain, together with high levels of everyday pain and dysfunction. Chapter 7 compared 40 OA and 40 matched healthy subjects. Chapter 8 evaluated 80 participants with knee OA. Chapter 9 applied a randomized, blinded, placebo-controlled design to evaluate the effects of a 14-day course of the Cox-2 inhibitor etoricoxib on ADI, QST and self-report neuropathic pain measures in order to investigate whether these proposed measures of central pain
augmentation are influenced by a largely peripherally-acting intervention. The sensitivity to change of the ADI was compared with that for CPT. Finally, a sub-analysis compared the responses of those with and without ADI-cold hyperalgesia.

**Results**

Studies 1 and 2 established that topical menthol evokes concentration-dependent sensory effects in healthy subjects, with higher concentrations producing significantly higher levels of intensity for each VAS sensation, significantly earlier response onset and higher PRI descriptor scores. Higher menthol concentrations consistently evoked more noxious descriptors such as icy, burning or stinging. Study 1 also found a significant association between CPT temperature and menthol response. When subjects were grouped dichotomously according to CPT <15°C, those with high CPT values recorded higher intensity of cold and unpleasantness and selected more noxious words during the menthol test. There were strong similarities between responses to equivalent concentrations of the liquid and gel formulations. Study 2 found that the higher of the two gel concentrations provided the better discrimination between normal and atypical responses across both intensity and quality measurement parameters. An atypical response in pain-free individuals was characterized by report of mild unpleasantness or pain associated with words such as icy, burning or stinging. Study 3 showed that the new ADI scoring system had good ability to discriminate between sustained thermal-cold temperatures and good internal reliability. Word choice at the lowest temperature closely matched that for the higher menthol concentration of Study 2. ROC curve analysis showed that an ADI cut-off score of 4.9 during sustained cold was able to predict CPT <15°C group with high sensitivity and specificity. ADI ≥5 was therefore selected as the ADI-cold hyperalgesia cut-off for future studies. Study 4 assessed test-retest reliability over 2 occasions in healthy subjects and found high ICC values for all aspects of the ADI. VAS intensity values showed greater variability between sessions than the descriptor quality sub-score, with all scores slightly lower on the second occasion.

Studies 5, 6 and 7 evaluated the menthol ADI as a test of cold hyperalgesia in an older clinical population. Study 5 compared OA and matched healthy older subjects and found that those with painful knee OA had significantly higher ADI scores, reporting higher unpleasantness and pain intensity and selecting noxious words such as burning and stinging more frequently. Study 6 investigated in more detail the association between ADI-cold hyperalgesia and additional QST and self-report measures of centrally augmented pain. 24% subjects were identified a cold hyperalgesic by both ADI ≥5 and CPT >15°C. Those with ADI ≥5 showed significantly greater cold, mechanical and heat hyperalgesia at both local and distant test sites and reported
higher incidence of significant neuropathic-type pain symptoms (PainDETECT). The high ADI group also reported significantly higher levels of pain and dysfunction and exhibited decreased function during physical tests. Study 7 found that the Cox-2 inhibitor etoricoxib was more effective than Placebo in improving pain and function and reducing local hyperalgesia. There were also some signs of influence over more centrally-controlled pain measures following just 14-days of etoricoxib, with forearm menthol-ADI score and PainDETECT score showing significant improvements. Although CPT temperature showed no change for those taking active etoricoxib, analysis of descriptors also selected at CPT in this study showed an 11% reduction in burning quality, very similar to that found for menthol. Overall however, the ADI showed considerably superior sensitivity to change than CPT. A sub-analysis of etoricoxib effect in high and low ADI groups showed that non-hyperalgesic subjects responded slightly better than those with higher ADI.

Conclusion
This investigation into the use of topical menthol as a test for cold hyperalgesia found that the particular gel formulation and ADI scoring system developed for these studies showed good levels of construct and criterion validity and good reliability between two test sessions. The close similarities in sensation quality experienced during graded menthol and sustained thermal cold stimuli suggest similar transduction and transmission mechanisms. The menthol ADI test was able to discriminate as sensitively as conventional CPT between normal and abnormal pain responses in both healthy and clinical cohorts. Comparison of reliability between CPT and menthol suggests that menthol ADI score may be a more reliable measure, although this needs further assessment in larger cohorts. However, given its demonstrated validity, the menthol test is potentially a more clinically applicable test for cold hyperalgesia due to its simplicity, low cost and relative ease of interpretation. The final study with a clinical population demonstrated the potential value of a measurement system that combines several pain domains (intensity and quality) and so is more sensitive than CPT to small changes in the pain response system. If the menthol test is applied at a site distant from the inflamed or injured area it has the potential to be able to identify those whose pain has started to develop centrally-augmented characteristics such as changes to response sensation quality or intensity. In addition, the ADI has sufficient sensitivity to signal whether an intervention is successfully changing either intensity or quality of an individual’s widespread pain response. A number of further studies are clearly needed to clarify these findings, but this investigation has shown that the menthol ADI test is a valuable additional QST tool that will enable greater investigation of cold hyperalgesia as a phenomenon associated with chronic pain.
Relevant Publications


Wright A, Benson H, Will R, Moss P (2012): Subjects with knee osteoarthritis exhibit widespread cold and mechanical hyperalgesia, associated with functional disability and lower quality of life compared with matched controls. 14th World Congress on Pain, Milan, Italy
Patents
Moss P, Wright A, Benson H.

• European Patent Application: 09806234.2 Method and Device for Determining the Severity of a Pain Disorder
• Australian Patent Application: Method and Device for Determining the Severity of a Pain Disorder
• United States Patent Application 13/057,836: Method and Device for Determining the Severity of a Pain Disorder
# Table of Contents

Declaration ........................................................................................................................................ iii  

Acknowledgements .......................................................................................................................... v  

Abstract .......................................................................................................................................... vii  

Relevant publications ...................................................................................................................... xi  

Table of contents ............................................................................................................................. xiii  

Abbreviations .................................................................................................................................. xxv  

Chapter 1  

Introduction ....................................................................................................................................... 1  

Chapter 2  

Literature Review ............................................................................................................................ 9  

2.1 Introduction ................................................................................................................................ 10  

2.2 Pain definitions ........................................................................................................................... 12  

2.3 Pain measures ............................................................................................................................ 15  

2.4 Abnormal pain processing in chronic pain ................................................................................ 25  

2.5 Cold hyperalgesia as an indicator of centrally-augmented pain ............................................. 51  

2.6 Topical menthol as a test for cold hyperalgesia ........................................................................ 67  

2.7 Conclusions ................................................................................................................................ 74  

Chapter 3  

Study 1 - Evaluation of the liquid test formulation  

3.1 Abstract ....................................................................................................................................... 75  

3.2 Background ............................................................................................................................... 77  

3.3 Hypotheses .................................................................................................................................. 81  

3.4 Method ....................................................................................................................................... 82  

3.5 Results ....................................................................................................................................... 88  

3.6 Discussion ................................................................................................................................... 101  

3.7 Summary .................................................................................................................................... 107  

Chapter 4  

Study 2 - Evaluation of the gel test formulation  

4.1 Abstract ....................................................................................................................................... 109  

4.2 Background .................................................................................................................................. 110  

4.3 Hypotheses .................................................................................................................................. 113  

4.4 Method ....................................................................................................................................... 114  

4.5 Results ....................................................................................................................................... 119
Chapter 5

Study 3 - Validation of the ADI scoring system using cold thermal stimuli

5.1 Abstract ..................................................................................................................................141
5.2 Background ............................................................................................................................143
5.3 Hypotheses ..............................................................................................................................146
5.4 Method ......................................................................................................................................147
5.5 Results ......................................................................................................................................153
5.6 Discussion ................................................................................................................................163
5.7 Summary ..................................................................................................................................170

Chapter 6

Study 4 – Evaluation of test retest reliability

6.1 Abstract ..................................................................................................................................171
6.2 Background ............................................................................................................................173
6.3 Hypotheses ..............................................................................................................................177
6.4 Method ......................................................................................................................................178
6.5 Results ......................................................................................................................................183
6.6 Discussion ................................................................................................................................191
6.7 Summary ..................................................................................................................................197

Chapter 7

Study 5 - Matched comparison between healthy subjects and those with knee osteoarthritis

7.1 Abstract ..................................................................................................................................199
7.2 Background ............................................................................................................................201
7.3 Hypotheses ..............................................................................................................................203
7.4 Method ......................................................................................................................................204
7.5 Results ......................................................................................................................................212
7.6 Discussion ................................................................................................................................224
7.7 Summary ..................................................................................................................................232

Chapter 8

Study 6 - Comparison between knee osteoarthritis subjects with high and low ADI scores

8.1 Abstract ..................................................................................................................................233
8.2 Background ............................................................................................................................235
8.3 Hypotheses ..............................................................................................................................237
8.4 Method ......................................................................................................................................238
8.5 Results ......................................................................................................................................245
8.6 Discussion ................................................................................................................................256
Chapter 9

Study 7 - The effect of etoricoxib on hyperalgesia, pain and function in subjects with knee OA

9.1 Abstract ......................................................................................................................265
9.2 Background ...............................................................................................................267
9.3 Hypotheses ...............................................................................................................270
9.4 Method ..................................................................................................................271
9.5 Results ..................................................................................................................278
9.6 Discussion .............................................................................................................297
9.7 Summary ...............................................................................................................307

Chapter 10

Discussion .....................................................................................................................309

Chapter 11

Summary .......................................................................................................................327

Chapter 12

References ....................................................................................................................331
Appendices ..................................................................................................................363
List of Figures

Chapter 2
Figure 2.1: Percentage of subjects with CPT >15°C, using datasets from a variety of
musculoskeletal disorders associated with chronic pain and poor resolution...64

Chapter 3
Figure 3.1: Differences in maximum VAS intensity ratings between menthol
concentrations (Liquids A, B and C). ..................................................89
Figure 3.2: Percentage of subjects reporting VAS >0 for each sensation during
application of each menthol concentration, and average maximum VAS intensity
for each sensation...........................................................................90
Figure 3.3: Mean quality descriptor index scores for each menthol concentration ....91
Figure 3.4: Percentage of subjects choosing the most commonly selected words for each
liquid. ..........................................................................................91
Figure 3.5: Percentage use of words types during 15-minutes application of each
concentration ...............................................................................92
Figure 3.6a-b: Correlations between PRI score at: a) Liquids B & C; b) Liquids A & B ...92
Figure 3.7: Time to onset of first recorded sensation for each concentration ..........93
Figure 3.8: Time to onset of first recorded sensation for each concentration ..........93
Figure 3.9: Time to peak (maximum) VAS of each sensation at each concentration.....94
Figure 3.10a-c: Mean VAS intensity at each minute during test time for each sensation
at each concentration: a) cold; b) unpleasantness; c) pain ..........................94
Figure 3.11a: Time-course for percentage of subjects selecting key words during Liquid
C.................................................................................................95
Figure 3.11b: Time-course for percentage of subjects selecting key words during Liquid
A .................................................................................................95
Figure 3.12: Differences in maximum VAS intensity ratings for subjects in high and low
CPT groups ..............................................................................98
Figure 3.13: Differences in PRI scores for subjects in high and low CPT groups ......98
Figure 3.14a-c: Comparison of most frequently chosen descriptors between those with
high or low CPT values for all three concentrations..................................99
Figure 3.15: Comparison between CPT < or >15°C groups for time to onset of VAS cold,
unpleasantness and pain at each concentration ......................................99

Chapter 4
Figure 4.1: Differences in maximum VAS intensity ratings between Gel A and Gel B .119
Figure 4.2: Percentage of subjects reporting at least some of each VAS sensation (VAS
>0) plus average maximum VAS intensity reported for each sensation at each
concentration ...........................................................................120
Figure 4.3: Overlaid scatterplots illustrating positive associations between Gels A and B for cold, heat, unpleasantness and pain ratings ................................................................. 120
Figure 4.4: Mean quality descriptor index scores for each menthol concentration....... 121
Figure 4.5: Percentage of subjects choosing the most commonly selected words for each gel ................................................................................................................. 121
Figure 4.6: Overlaid scatterplots illustrating positive associations between Gels A and B for PRI and NWC scores .............................................................................. 122
Figure 4.7: Time to onset of first report for each VAS sensation at each concentration ................................................................................................................. 122
Figure 4.8: Time to maximum VAS for each sensation and each concentration .......... 123
Figure 4.9a-d: Mean VAS intensity at each minute during test time for each sensation at each concentration: a) cold VAS; b) heat VAS; c) unpleasantness VAS; d) pain. ................................................................................................................................. 123
Figure 4.10a: Time-course for selection of key words during Gel B: percentage of subjects selecting each word every 2 minutes ......................................................... 124
Figure 4.10b: Time-course for selection of key words during Gel A.......................... 124
Figure 4.11a: Maximum VAS ratings for cold, unpleasantness and pain for equivalent Liquid and Gel formulations .................................................................................. 125
Figure 4.11b: PRI scores for equivalent Liquid and Gel formulations .......................... 125
Figure 4.12: Percentage choice of key words for equivalent B concentration Liquid and Gel ............................................................................................................. 125
Figure 4.13a: Time to onset for each VAS sensation for equivalent Liquid and Gel formulations ............................................................................................................ 126
Figure 4.13b: Time to peak for each VAS sensation for equivalent Liquid and Gel formulations ............................................................................................................. 126
Figure 4.14a: High and low K-means cluster groups for Gel A, comparing VAS ratings. ......................................................................................................................... 127
Figure 4.14b: High and low K-means cluster groups for Gel B, comparing VAS ratings. ......................................................................................................................... 127
Figure 4.15: Percentage of subjects in high and low cluster groups selecting key words ............................................................................................................. 128
Figure 4.16: Percentage of subjects in the high score cluster for each concentration. Horizontal lines designate a proposed target zone for a cold hyperalgesic group between 15 and 25% of a healthy subjects ................................................. 130

Chapter 5
Figure 5.1: Electronic VAS potentiometer box with sliders for each sensation of cold, heat, unpleasantness and pain ................................................................. 150
Figure 5.2: Histograms showing normal distribution curves for ADI during sustained cold at 10°C and 15°C ......................................................................................... 153
Figure 5.3: ADI and MWS sub-score values for each sustained cold temperature ....... 154
Figure 5.4: VAS intensity values for cold, heat, unpleasantness and pain at each sustained temperature (10°C, 15°C and 20°C) .................................................................154
Figure 5.5: Percentage of subjects choosing the most commonly selected words for each sustained temperature ..........................................................................................155
Figure 5.6a-d: Release patterns at each temperature for a) cold, b) heat, c) unpleasantness, d) pain. ......................................................................................................................155
Figure 5.7: Overlay scatterplot showing the correlation between cold pain threshold and MWS descriptor sub-score at 10°C and at 15°C ..............................................158
Figure 5.8: Differences in VAS ratings for cold, heat, unpleasant and pain at each sustained cold temperature between CPT groups ....................................................................159
Figure 5.9: Percentage choice of key descriptors during 10°C sustained cold and menthol Gel B ..................................................................................................................160
Figure 5.10: Comparison between sustained 10°C cold and Gel B (Study 2) for percentage of subjects reporting any cold, heat, unpleasantness or pain (primary axis) and maximum VAS rating for each sensation (secondary axis) ...........................................................................................................160
Figure 5.11a-b: Release patterns during application of: a) sustained 10°C cold, b) menthol Gel B ......................................................................................................................161

Chapter 6
Figure 6.1: Histograms showing normal distribution curves for menthol ADI data on Day 1 and Day 2 test sessions ................................................................................... 183
Figure 6.2a: Mean (SEM) values for total ADI score and MWS descriptor sub-score for Day 1 and Day 2 test sessions .................................................................................. 184
Figure 6.2b: Overlay scatterplot for ADI scores for Day 1 and Day 2 (yellow) and MWS scores for Day 1 and Day 2 (green) .............................................................................. 185
Figure 6.3: Percentage of subjects selecting specific words on Day 1 and Day 2 ......... 185
Figure 6.4: Percentage selection of different types of words on each test occasion ..... 186
Figure 6.5a: Percentage of subjects rating VAS >0 on Days 1 and 2 for cold, heat, unpleasantness and pain ............................................................................................. 186
Figure 6.5b: Mean (SEM) VAS ratings for those subjects scoring VAS >0 on test Days 1 and 2 for cold, heat, unpleasantness and pain ..................................................................... 186
Figure 6.6a: Overlay scatterplot for cold VAS ratings (blue) for Day 1 and Day 2 and heat VAS ratings (red) for Day 1 and Day 2 ........................................................................... 187
Figure 6.6b: Overlay scatterplot for unpleasantness VAS ratings (green) for Day 1 and Day 2 and pain VAS ratings (orange) for Day 1 and Day 2 ................................................. 187
Figure 6.7a-b: Intra-class correlation coefficients Day 1 to 2 for mean times to a) onset and b) peak VAS sensations .......................................................................................... 188
Figure 6.8a-b: Time course of response to menthol, showing mean VAS cold, heat, unpleasantness and pain at each minute during application for: a) Day 1 and b) Day 2 ............................................................................................................................. 189
Figure 6.9: Mean (SEM) VAS ratings for cold, heat, unpleasantness and pain on Day 1 and Day 2, comparing male and female subjects .......................................................... 190
Figure 6.10: Percentage of subjects reporting different descriptor words at CPT, comparing three separate test sessions .......................................................... 193

Chapter 7

Figure 7.1: Histogram showing the normal distribution curve for menthol ADI data across all subjects ........................................................................................................ 213
Figure 7.2a-b: Mean time to onset and time to peak for each VAS sensation ................. 214
Figure 7.3a-b: Percentage of subjects reporting some VAS intensity (>0/100) for each sensation (bar chart, axis 1); Average maximum intensity (0-100) for those reporting VAS>0 for each sensation (line graph, axis 2) ........................................ 216
Figure 7.4: OA subjects: mean time to onset and time to peak for each VAS sensation .......................................................................................................................... 216
Figure 7.5a-b: Mean VAS intensity at each minute for each sensation ............................ 217
Figure 7.6: Percentage choice of key descriptor words by healthy and OA subjects... 217
Figure 7.7: Mean pressure pain thresholds at each test site, comparing subjects divided according to CPT< or >15°C or ADI< or ≥5 ......................................................... 219
Figure 7.8: Mean times for each ALF task, comparing subjects divided according to CPT< or >15°C or ADI< or ≥5 ................................................................................. 219
Figure 7.9: Mean cold pain thresholds at the OA knee, unaffected knee (or equivalents) and elbow for OA and healthy control subjects ......................................................... 220
Figure 7.10: Mean pressure pain threshold at the OA knee, unaffected knee (or equivalents) and elbow for OA and healthy control subjects ................................................. 220
Figure 7.11: Mean heat pain threshold at the OA knee, unaffected knee (or equivalents) and elbow for OA and healthy control subjects ......................................................... 221
Figure 7.12: Mean time taken to complete chair transfer, 8-metre walk and stairs tasks for OA and healthy control subjects ......................................................... 222

Chapter 8

Figure 8.1: Algotect Descriptor Score (ADI) scoring structure ..................................... 242
Figure 8.2: Histogram showing the normal distribution curve for ADI data from all subjects with knee osteoarthritis ........................................................................... 246
Figure 8.3: Percentage of subjects selecting highest ranked words to describe sensations during the menthol stimulus and at CPT (mean word choice for the three sites). ........................................................................................................... 247
Figure 8.4: ADI and PainDETECT scores for subjects grouped according to sites exhibiting CPT>15°C moved to after 8.2 ........................................................................ 248
Figure 8.5: Mean scores for low and high ADI groups for WOMAC pain, stiffness and function sub-scores ................................................................................................. 249
Figure 8.6: Time taken (secs, secondary axis) and pain experienced (VAS, primary axis) during the three ALF tasks, for subjects with menthol ADI score <5 and ≥5....250
Figure 8.7: Mean cold pain threshold for low (<5) and high (≥5) ADI groups at each test site. ........................................................................................................................................251
Figure 8.8a-b: Most frequently chosen words at CPT for each test site: a) Non cold hyperalgesic subjects (forearm ADI <5); b) Cold hyperalgesic subjects (forearm ADI ≥5). ........................................................................................................................................251
Figure 8.9: Mean pressure pain threshold for low (<5) and high (≥5) ADI groups at each test site. ........................................................................................................................................252
Figure 8.10: Mean heat pain threshold for low and high ADI groups at each test site. ...252
Figure 8.11: Mean pain quality questionnaire scores for low and high ADI groups: PainDETECT and PQAS paroxysmal, surface and deep sub-scores. ..................253

Chapter 9
Figure 9.1: Etoricoxib intervention study protocol ..........................................................273
Figure 9.2: Mean WOMAC-pain scores for Placebo and Active groups at each time point ........................................................................................................................................279
Figure 9.3: Mean WOMAC-function scores for Placebo and Active groups at each time point ........................................................................................................................................280
Figure 9.4: VAS pain rating before, during and after ALF chair, walk and stairs tasks on Day 0 and Day 14: a) Active group; b) Placebo group ........................................281
Figure 9.5: Mean (SEM) cold pain threshold values for Placebo and Active groups at each time point for the OA knee .................................................................................................282
Figure 9.6: Mean (SEM) ADI values for Placebo and Active groups at each time point ........................................................................................................................................283
Figure 9.7a-b: Mean (SEM) cold pain threshold values for Placebo and Active groups at each time point for a) unaffected knee; b) elbow ........................................................................284
Figure 9.8: Mean (SEM) PainDETECT values for Placebo and Active groups at each time point ........................................................................................................................................286
Figure 9.9: Mean (SEM) PQAS sub-scores for Placebo and Active groups at each time point: paroxysmal, surface and deep sub-scores ....................................................287
Figure 9.10a-h: Mean (SEM) Active and Placebo group pain VAS scores at Day 0, 4 and 14 for low and high ADI groups ......................................................................................289-290
Figure 9.11a-c: Mean (SEM) for ALF physical function tasks at Day 0, 4 and 14 for low and high ADI groups: a) chair task; b) walk task; c) stairs task ........................................290
Figure 9.12a-b: Mean (SEM) PainDETECT scores at Day 0, 4 and 14 for low and high ADI groups: a) ADI<5 group; b) ADI≥5 group ...............................................................291
Figure 9.13a): Mean (SEM) PQAS ‘paroxysmal’, ‘surface’ and ‘deep’ sub-scores at Day 0, 4 and 14 for the low ADI group ..................................................................................292
Figure 9.13b): Mean (SEM) PQAS ‘paroxysmal’, ‘surface’ and ‘deep’ sub-scores at Day 0, 4 and 14 for the high ADI group ..................................................................................292
Figure 9.14a-b: Mean (SEM) menthol ADI scores at Day 0, 4 and 14 for a) the low ADI group and b) the high ADI group ................................................................. 293
Figure 9.15a): Mean (SEM) pressure pain threshold values at Day 0, 4 and 14 for the low ADI group for OA knee, unaffected knee and elbow test sites .................. 294
Figure 9.15b): Mean (SEM) pressure pain threshold values at Day 0, 4 and 14 for the high ADI group for OA knee, unaffected knee and elbow test sites .................. 294

List of Tables

Chapter 3
Table 3.1: Inclusion / Exclusion criteria for liquid formulation study ................................. 82
Table 3.2: Age characteristics between male and female subjects ................................. 88
Table 3.3: Mean (SEM) area under the time-VAS curve values for each sensation at each concentration .................................................................................................................. 89
Table 3.4: During Liquid C, subjects in the high PRI score group (score > mean +1SD) reported significantly higher VAS intensity for cold and unpleasantness, but not pain ................................................................. 96
Table 3.5: Correlations between CPT (°C) and PRI and NWC descriptor indices at each concentration .......................................................................................................................... 97

Chapter 4
Table 4.1: Inclusion / Exclusion criteria for liquid formulation study............................... 114
Table 4.2: Number (percentage) of subjects classified in high or low VAS cluster groups following dichotomous K-means cluster analysis ........................................ 126
Table 4.3: Individual membership of high cluster groups for each VAS sensation at each concentration of gel (✓). Mean VAS given when individuals did not meet the VAS cut-off value for high cluster membership ....................................................... 127
Table 4.4: Number (percentage) of subjects classified in high or low PRI cluster groups following dichotomous K-means cluster analysis ................................. 128
Table 4.5: Classification of subjects as cold hyperalgesic or not according to membership of high VAS cluster groups, high PRI cluster groups and choice of key words. VAS / PRI score > than the cut-off is designated ✓ VAS values < the cut off are recorded .......................................................................................................................... 129
Table 4.6: Membership of each VAS and PRI high cluster group (with cut-off values shown) and identification of choice of key words icy, burning and stinging, VAS values shown when cut-off value not reached ................................................................. 130

Chapter 5
Table 5.1: Inclusion / Exclusion criteria for ADI validation study ........................................... 147
Table 5.2: Correlations between temperatures for each VAS intensity ............................. 156
Table 5.3: Correlations between PRI score and ADI total and MWS sub-score at each sustained temperature ...............................................................................................................157
Table 5.4: Best-fit ROC curve data for sensitivity and specificity of PRI and ADI indices to predict membership of the CPT >15°C group .................................................................157
Table 5.5: ADI total and MWS descriptor sub-scores for each CPT group (< or >15°C) at each sustained temperature ...........................................................................................................158
Table 5.6: Mean (SD) and 95% confidence interval data for ADI and MWS sub-score values for 10°C sustained cold and Gel B (Study 2). .................................................................159

Chapter 6
Table 6.1: Inclusion / Exclusion criteria for liquid formulation study ...........................................178
Table 6.2: Age characteristics of male and female subjects ..........................................................183
Table 6.3: Intra-Class Correlation Coefficient (95% confidence interval) for ADI total score and MWS and VAS sub-scores ..........................................................................................184
Table 6.4: Means (SD) for male and female ADI, MWS and VAS scores on Day 1 and Day 2 ...........................................................................................................................................189

Chapter 7
Table 7.1: Inclusion / Exclusion criteria for subjects with knee OA ............................................204
Table 7.2: Inclusion / Exclusion criteria for Healthy control subjects ..........................................205
Table 7.3: Subjects with knee OA and controls were matched by 9-year age band and gender .........................................................................................................................................205
Table 7.4: Standardised knee and elbow test sites for QST tests .....................................................206
Table 7.5: Comparisons between OA and matched control subjects for BMI and quality of life ........................................................................................................................................212
Table 7.6: Correlations between ADI and MWS scores and CPT temperature at the OA knee, unaffected knee and elbow .........................................................................................215
Table 7.7: Cross-tabulation of membership of high and low ADI and CPT groups by OA and healthy subjects ........................................................................................................218
Table 7.8: Cold, heat and punctate detection thresholds for OA and healthy control subjects at each test site ..............................................................................................................222
Table 7.9: SF-36 physical and mental health sub-scores for OA and healthy control subjects .................................................................................................................................223
Table 7.10: Means, standard deviation, 95% confidence interval and variance for CPT values (°C) at the OA knee and elbow for OA and healthy subjects ........................................231

Chapter 8
Table 8.1: Inclusion / Exclusion criteria ......................................................................................238
Table 8.2: Standardised knee and elbow test sites for QST tests ...................................................240
Table 8.3: Comparisons between OA subjects with ADI < or ≥5 for gender, age and BMI .................................................................246
Table 8.4: Membership of high and low ADI groups (< or ≥5) compared with membership of high and low mean CPT groups (< or >15°C average all sites)247
Table 8.5: Percentage of subjects in each ADI group reporting pain >0/10 VAS during and after each ALF task..........................................................................................................................250
Table 8.6 Cross-tabulation between membership of PainDETECT categories and ADI categories ......................................................................................................................................................253
Table 8.7: Comparison of cold (CPT), heat (HPT) and pressure (PPT) pain thresholds between older healthy subjects from Study 5 and OA subjects from the current study........................................257

Chapter 9
Table 9.1: Etoricoxib study: inclusion and exclusion criteria ........................................271
Table 9.2: Standardised knee and elbow test sites for QST tests .......................................273
Table 9.3: Baseline characteristics of Placebo and Active groups.....................................278
Table 9.4: Mean (SEM) time taken (secs) to complete Aggregated Locomotion Function tasks for Placebo and Active groups at Day 0, Day 4 and Day 14.................................280
Table 9.5: Change in ADI score components: mean (SEM) values for sensation quality (MWS) and intensity (VAS cold, heat, unpleasantness and pain) for Active and Placebo groups at Days 0, 4 and 14.................................................................283
Table 9.6: Mean (SEM) values pressure pain thresholds at the OA knee, unaffected knee and elbow test sites for Active and Placebo groups on Days 0, 4 and 14........285
Table 9.7: Mean (SEM) values heat pain thresholds at the OA knee, unaffected knee and elbow test sites for Active and Placebo groups on Days 0, 4 and 14.............285
Table 9.8: Change in PainDETECT category (‘negative neuropathic’ <13; ‘unclear neuropathic’ 13-18; ‘positive neuropathic’ ≥19) between Day 0 and Day 14 for Placebo and Active groups.................................................................285
Table 9.9: Comparison of different measures of responsiveness to etoricoxib treatment for ADI, PainDETECT, cold pain threshold and pressure pain threshold...........288
Table 9.10: Numbers of subjects with ADI < or ≥5 allocated to Placebo or Active groups ...........................................................................................................................................................................289
Table 9.11: PainDETECT group change by ADI group, Day 0 to Day 14.........................291
Table 9.12: Mean(SEM) cold pain threshold values at Day 0, 4 and 14 for low and high ADI groups .................................................................................................................................................................293
Table 9.13: Comparison between ADI groups of effect size and Cohen’s d for key study outcome measures .................................................................................................................................295
Table 9.14: Difference in CPT, ADI and elbow PPT values between those categorised as ‘positive’ or ‘negative’ neuropathic for PainDETECT at Day 14 ........................................303
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADI</td>
<td>Algometrographic Index (menthol test index score)</td>
</tr>
<tr>
<td>ALF</td>
<td>Aggregated locomotion function score</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>CDT</td>
<td>Cold detection threshold</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CPM</td>
<td>Conditioned pain modulation</td>
</tr>
<tr>
<td>CPT</td>
<td>Cold pain threshold</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>DN4</td>
<td>Doleur Neuropathique 4</td>
</tr>
<tr>
<td>DNIC</td>
<td>Diffuse noxious inhibitory control</td>
</tr>
<tr>
<td>e-VAS</td>
<td>Electronic VAS</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional magnetic resonance imaging</td>
</tr>
<tr>
<td>GC</td>
<td>Gas chromatography</td>
</tr>
<tr>
<td>HDT</td>
<td>Heat detection threshold</td>
</tr>
<tr>
<td>HPC</td>
<td>Hydroxypropyl cellulose</td>
</tr>
<tr>
<td>HPT</td>
<td>Heat pain threshold</td>
</tr>
<tr>
<td>ILC</td>
<td>Interleukin C</td>
</tr>
<tr>
<td>LANSS</td>
<td>Leeds Assessment of Neuropathic Symptoms and Signs</td>
</tr>
<tr>
<td>LTP</td>
<td>Long term potentiation</td>
</tr>
<tr>
<td>MCS</td>
<td>Mental health category score (a sub-score of the SF-36 survey)</td>
</tr>
<tr>
<td>MPQ</td>
<td>McGill Pain Questionnaire</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>MWS</td>
<td>Mean Word Score (descriptor sub-score of the ADI index)</td>
</tr>
<tr>
<td>NPQ</td>
<td>Neuropathic Pain Questionnaire</td>
</tr>
<tr>
<td>NRS</td>
<td>Numerical rating scale</td>
</tr>
<tr>
<td>NSAID</td>
<td>Non-steroidal anti-inflammatory</td>
</tr>
<tr>
<td>NWC</td>
<td>Number of words chosen (from McGill)</td>
</tr>
<tr>
<td>OA</td>
<td>Osteoarthritis</td>
</tr>
<tr>
<td>PAG</td>
<td>Periaqueductal grey</td>
</tr>
<tr>
<td>PdDT</td>
<td>Punctate detection threshold</td>
</tr>
<tr>
<td>PdT</td>
<td>Punctate pain threshold</td>
</tr>
<tr>
<td>PCS</td>
<td>Physical category score (a sub-score of the SF-36 survey)</td>
</tr>
<tr>
<td>PGE2</td>
<td>Prostaglandin E2</td>
</tr>
<tr>
<td>PPT</td>
<td>Pressure pain threshold</td>
</tr>
<tr>
<td>PQAS</td>
<td>Pain Quality Assessment Scale</td>
</tr>
<tr>
<td>PRI</td>
<td>Pain Rating Index (from McGill)</td>
</tr>
<tr>
<td>QoL</td>
<td>Quality of life</td>
</tr>
<tr>
<td>QST</td>
<td>Quantitative sensory testing</td>
</tr>
<tr>
<td>ROC curve</td>
<td>Receiver operating characteristic curve</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error of the mean</td>
</tr>
<tr>
<td>TRPA1</td>
<td>Transient receptor potential ankirin 1</td>
</tr>
<tr>
<td>TRPM8</td>
<td>Transient receptor potential melastatin 8</td>
</tr>
<tr>
<td>TRPV1</td>
<td>Transient receptor potential vanilloid 1</td>
</tr>
<tr>
<td>TTO</td>
<td>Time to onset (for VAS intensity)</td>
</tr>
<tr>
<td>TTP</td>
<td>Time to peak (for VAS intensity)</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual analogue scale</td>
</tr>
<tr>
<td>WOMAC</td>
<td>Western Ontario and McMaster Universities Osteoarthritis Index</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The overall purpose of this investigation was to develop a new, clinically relevant test for cold hyperalgesia that would help to identify individuals with features of centrally-augmented chronic pain.

Background

Chronic pain is a widespread and costly world-wide problem, associated with high levels of dysfunction, poor quality of life and a heavy economic burden. It has been estimated that one in five adults suffers from ongoing moderate to severe pain that has lasted more than 6 months (Breivik et al., 2006). Chronic pain creates an added burden for health and welfare systems, estimated at an annual minimum of $635 billion in the US, and further billions of dollars in lost productivity (Gaskin & Richard, 2012). Although chronic pain is often associated with specific disorders such as low back pain or fibromyalgia, anecdotal evidence suggests that non-resolving pain can potentially develop from a wide range of conditions. Chronic pain is generally diagnosed by longevity of symptoms and failure to respond to escalating interventions. The consequent development of secondary effects, whether psychological or physiological, may reinforce the chronic pain state. Established chronic pain states can often be difficult to change, with patients not responding successfully to analgesics or other pain treatments (Conway & Higgins, 2011). Earlier identification of individuals who are likely to develop chronic pain would enable more individually targeted treatment to be provided, before secondary effects had become entrenched. The personal and economic benefits could potentially be wide-ranging.

There is increasing evidence that development of chronic pain is associated with signs of abnormal augmentation in central pain processing. Studies of chronic pain conditions such as fibromyalgia have shown that chronicity and severity are associated with a wide range of physiological, psychological and cognitive changes. Widespread hyperalgesia, changes in endogenous pain mechanisms, neuropathic-type pain symptoms and altered mood and cognition have been reported (Kindler et al, 2011). In other chronic pain conditions such as whiplash or osteoarthritis (OA) similar widespread hyperalgesia, dysfunctional descending pain inhibition and presence of neuropathic-type pain in the absence of neuropathy have also been demonstrated in
those reporting more severe pain and dysfunction (Arendt-Nielsen et al., 2010; Hochman et al., 2011). This has led to the suggestion that for some individuals, patho-
etiology has become irrelevant and self-perpetuating pain has become the driver. (Gwilym et al., 2008).

There are many proposed methods of testing for central pain augmentation, although few appear appropriate for easy translation into the clinical context where they are needed. Self-report tools are popular due to their easy application. Questionnaires that evaluate extent of neuropathic-type symptoms have been proposed as a simple clinical approach to differentiating those with spontaneous symptoms that may reflect altered pain processing. Good evidence is emerging that questionnaires such as PainDETECT (Freynhagen et al., 2006) or the Pain Quality Assessment Scale (Victor et al., 2008) may be useful in identifying patients whose pain is no longer clearly driven by peripheral nociceptive processes (Gwilym et al., 2009). However, there is disagreement about correct interpretation of these tests (Gauffin et al., 2013), since a ‘positive’ score may reflect undiagnosed neuropathy rather than central pain augmentation (Baron et al., 2010). Quantitative sensory tests (QST) are widely used in a research context to demonstrate the characteristics and potential mechanisms of changes in pain processing. Widespread mechanical hyperalgesia, temporal summation and diffuse noxious inhibitory control / central pain modulation assessments have been used to illustrate in vivo central sensitisation and altered inhibitory control effects (Berglund et al., 2002; Leffler et al., 2002; Staud et al., 2007). However, it is unclear which is the most valid indicator of centrally-augmented pain and reliable application of these tests in the clinical environment is problematic. Brain imaging techniques may one day be the ‘gold standard’ of pain assessment tools. However, currently there does not appear to be a consensus on the most reliable method for differentiating the imaging characteristics of abnormally sensitised pain (Tracey, 2005; Lee et al., 2008; Tracey, 2010).

The phenomenon of cold hyperalgesia has been suggested as a potential key marker for individuals with centrally-augmented pain. Significantly elevated cold pain thresholds (CPT) have been reported across a range of poorly resolving central, neurogenic and musculoskeletal chronic pain disorders (Piche et al., 2010; Zanette et al., 2010; Coombes BK et al., 2012; Goldsmith et al., 2012), always associated with higher levels of pain and dysfunction (Hurtig et al., 2001; Sterling et al., 2006). Cold hyperalgesia is reported to be a key characteristic of those with more severe presentations in central sensitivity syndromes such as fibromyalgia or chronic regional pain syndrome.
(Desmeules et al., 2003; Vaneker et al., 2005). Elevated CPT and changes in the quality of sensation during cold pain have been shown to be closely associated with dysfunction in pain inhibitory systems (Yarnitsky et al., 2008; Piche et al., 2011) and it is rare to find heterogeneity in CPT across affected and unaffected body sites, implying a centrally controlled mechanism. Yet, crucially, CPT does not appear to be a fixed phenomenon and has shown the capacity to return to normal following appropriate intervention (Smith et al., 1999; Kosek & Ordeberg, 2000).

However, the current testing methods for cold hyperalgesia are problematic and impractical for clinical use. Cold pain threshold is most often tested using a peltier thermode. Aside from the expense and delicate nature of the equipment, peltier thermodes struggle to achieve normal range temperatures below 10°C. Since “hyperalgesia” is most commonly defined as a significant difference from the control group mean, an unreliable control mean is a problem. More recently, simpler, more clinically appropriate approaches to testing for abnormal cold response have been proposed, using application of ice (Maxwell & Sterling, 2012) or of an object such as a glass rod at a temperature of around 20°C (Baron et al., 2010). Both approaches succeed in terms of simplicity but fail to achieve good validity or reliability.

The current investigation therefore sought to develop a new test for cold hyperalgesia by applying basic science understanding of alternative cold receptor agonists. TRPM8 has been found to be the predominant transducer of cold sensation. As well as being activated by cold, TRPM8 is also activated by l-menthol (Peier et al., 2002). Although the mechanisms by which cold and menthol activate TRPM8 are different, their overall sensory effects appear to be similar. Steadily decreasing temperature is coded as increasingly noxious. Basic science research suggests that the TRPM8 channel is structured so that increased binding of menthol molecules steadily increase calcium influx (Janssens & Voets, 2011) and, although the concentration-dependent effects of menthol have not been systematically investigated, comparison of the sparse studies that have used different concentrations of menthol as a skin sensitisier suggest that higher concentrations evoke increasingly noxious responses (Green, 1992; Wasner et al., 2004). Since the phenomenon of cold hyperalgesia appears to be a centrally-mediated phenomenon, noxious response to innocuous cold temperature should theoretically be mirrored by noxious response to an innocuous concentration of menthol. However, this study sought to investigate this hypothesis.
Chapter 1

Introduction

All pain is a multi-dimensional phenomenon, a fact that is not necessarily reflected with individual QST test results. As with other conventional QST tests, cold pain threshold (CPT) assessment results in a single temperature value, which represents the phenomenon of response to cold. However, studies using sustained thermal cold stimuli have found that there are distinct changes in sensation quality as well as sensation intensity when noxious cold is experienced (Davis & Pope, 2002; Frolich et al., 2010). Quantification of cold pain by using conventional CPT methodology may therefore not be optimal. A test for cold pain that specifically evaluated the domains of sensation intensity and sensation quality / type would provide more comprehensive information about mechanisms driving the response and so about degree of hyperalgesia / allodynia. This information could potentially assist in decisions about most appropriate treatment for an individual with cold hyperalgesia. Therefore, alongside development of an optimal topical menthol formulation, the current investigation sought to develop a response measurement system that included the elements of sensation intensity and quality.

Validity and reliability

The menthol test is intended as an assessment that can be used for both research and clinical purposes. The ultimate value of the test is therefore determined by the extent to which it is able to demonstrate validity and reliability, alongside a simple application and interpretation method that can be applied in a clinical setting.

Validity involves the extent to which a test measures what it aims to (Portney & Watkins, 2008). Criterion, content and construct validity are primarily considered in this series of studies. Criterion validity establishes that the test is as good as a ‘gold standard’ test in measuring the construct. Content validity establishes whether the test sufficiently samples content that would be considered to be essential to the construct being measured and construct validity indicates that the test reflects the theoretical elements of the construct being tested. The starting point for validity testing therefore must be with a clear definition of the construct to be tested or compared. For cold hyperalgesia however, this is problematic and somewhat imprecise as there is no consensus on use of terminology and associated definitions. As discussed in Chapter 2.2, the terminology of cold hyperalgesia and cold allodynia are often used interchangeably, even though the IASP taxonomy advises that ‘allodynia’ should only be used to describe response to a stimulus modality that can only be innocuous, such as brushing (IASP, 2011 update). For a stimulus such as cold that can be both innocuous
and noxious the IASP recommendation is to use the term ‘hyperalgesia’. Therefore a noxious response to any cold temperature should be described as cold hyperalgesia.

Cold hyperalgesia is most often defined by response to the current ‘gold standard’ test for cold pain threshold (CPT) using a peltier thermode and method of limits. A cold hyperalgesic response is one where the selected temperature at which the individual feels painful cold is significantly higher than the normal mean CPT. So ‘cold hyperalgesia’ tends to be used as a relative rather than an absolute term. From the perspective of criterion validity this means that the menthol test will need to demonstrate concurrently a clear association with CPT such that individuals who exhibit high CPT also exhibit high menthol test result. This investigation will also use the definition of cold hyperalgesia as CPT temperature value >15°C as an alternative concurrent validity criterion. Although open to debate, the justification for use of this cut-off is provided in Section 2.5.

Content validity relates to whether a test includes all the content that would be considered to be essential to the construct being measured. Once again this is not clear-cut with cold hyperalgesia. Cold hyperalgesia is quantified during conventional CPT testing as a simple temperature value, with this value reflecting two elements: the target sensation (painful cold) and the temperature at which it is reported. If this temperature is higher than normal for a noxious response, this is defined as cold hyperalgesia. Yet studies that have evaluated cold response at sustained low temperatures suggest that the noxious response is often more complex and involves changes in sensation quality in combination with changes in intensity (Davis & Pope, 2002). The term ‘noxious’ may also encompass the less intense construct of unpleasantness as well as that of pain. Basic science studies have also demonstrated that noxious cold is not black and white. Cold at decreasing temperatures activates a range of nerve fibres, which are able to code a range of different sensation qualities. These may be noxious to a lesser or greater extent. For example, activation of cold and heat sensing low threshold Aδ-nociceptors may well be felt as a mildly unpleasant cold with stinging qualities whereas activation of polymodal c-fibres is likely to be felt as a burning sensation with no cold quality (McKemy, 2005). For ideal construct validity therefore, a cold hyperalgesic test needs to be able to identify changes both in intensity and sensation quality. It is possible therefore that the menthol test may be able to offer better content validity than the CPT test.
An unreliable test cannot be considered valid. The current investigation will therefore apply two types of reliability measure. Within Studies 1, 2 and 3, internal consistency of both the menthol stimulus and the test scoring system will be evaluated to see if a healthy individual tends to select similar types of sensation when rating their experience at difference concentrations of menthol and also when the scoring system is applied to different sustained cold temperatures. It is also important to assess whether a healthy individual’s responses to the same menthol stimulus are consistent over time. Test-retest reliability will therefore be assessed in Chapter 6.

Investigation aims and strategy
The overall purpose of this investigation was to develop a clinically-relevant test for cold hyperalgesia that would assist in identifying individuals with centrally-augmented pain. Following lab development of the menthol formulation a series of seven studies were conducted. The first four studies evaluated basic validity and reliability in healthy subjects. In the final three studies, the revised menthol test and Algotect Descriptor Index (ADI) scoring system was evaluated for discriminative ability in a clinical population with knee OA. OA is a chronic pain condition with high incidence and high heterogeneity. Knee OA in particular has been associated with increased pain and dysfunction and, for approximately 20% of individuals, persistent severe pain even after replacement of the joint (Wylde et al., 2011), associated with less severe pre-operative radiological signs (Valdes et al., 2012). Additional indications of centrally-augmented pain, such as mechanical hyperalgesia, temporal summation and reduced descending inhibitory pain control have also been reported in a proportion of those with more severe knee OA (Bajaj et al., 2001; Deutsch et al., 2001; Gwilym et al., 2009; Arendt-Nielsen et al., 2010; Goldenberg et al., 2011).

1. **Develop a menthol formulation that demonstrates ability to discriminate clearly between typical and atypical responses to cold**

The initial aim of this investigation was to develop and evaluate a menthol formulation that would be able to discriminate between different sensory responses to cold. Formulation development was carried out in a laboratory setting, combining pharmacological preparation, release experiments and small psychophysical pilot tests. This is described briefly in Appendix 1. Evaluation of the formulation at different concentrations was carried out in healthy pain-free subjects, first using a simple liquid solution and then a gel-based preparation. Dose-dependent effects and validity compared with CPT values were evaluated (Chapter 3). The aim was to determine the
optimal formulation and menthol concentration that would differentiate between a 
non-hyperalgesic (normal) and hyperalgesic (atypical) response in healthy subjects 
(Chapter 4).

2. **Develop a measurement system that enables reliable and valid discrimination 
between normal and abnormal responses to cold**

Alongside development of the formulation, a measurement method was devised which 
combined appropriately weighted intensity and quality scores into an index measure 
(Appendix 2). This scoring system was evaluated for construct and criterion validity 
using conventional thermal-cold stimuli (Chapter 5). Test-retest reliability was then 
assessed (Chapter 6).

3. **Evaluate the ability of the menthol test to identify centrally-augmented pain in 
a clinical population**

Following initial development in healthy subjects, the test was applied to a clinical 
population of subjects with knee OA. The initial aim was to assess the discriminative 
ability of the menthol test. Initial comparisons were made between those with OA knee 
pain and age and gender-matched healthy controls (Chapter 7). A larger OA knee pain 
cohort was then evaluated using a larger bank of QST, self-report and physical function 
tests to investigate whether menthol test result was associated with additional signs of 
centrally-augmented pain (Chapter 8).

4. **Apply the menthol test to assess the efficacy of a Cox-2 inhibitor anti-
  inflammatory intervention in reversing measures of centrally-augmented pain 
in knee OA**

The final study (Chapter 9) used a double-blind, placebo-controlled design to 
investigate the effect of a 14-day course of the Cox-2 anti-inflammatory etoricoxib on a 
range of self-report and QST measures reflecting local and widespread pain 
mechanisms. Change in forearm menthol test score was assessed alongside QST at the 
knee and elbow, neuropathic-type pain quality to spontaneous pain and self-report of 
everyday pain and dysfunction. The study also explored whether drug efficacy was 
reduced in those who exhibit cold hyperalgesia.
Chapter 2

Literature Review

Overview of the Literature Review

2.1 Introduction

2.2 Pain Definitions

2.3 Pain Measures

Introduction
Tools for assessing pain processing
Summary

2.4 Abnormal pain processing in chronic pain

Pain as the disorder
The development of abnormal pain processing
Conditions exhibiting signs of central pain augmentation
Summary

2.5 Cold hyperalgesia as an indicator of centrally-augmented pain

The incidence of cold hyperalgesia
Assessing cold hyperalgesia
A definition of cold hyperalgesia?
Summary

2.6 Topical menthol as a test for cold hyperalgesia

Current test methods
Menthol as an alternative test stimulus
Summary

2.7 Conclusions
2.1 Introduction
This series of studies aimed to develop an alternative quantitative sensory test for cold hyperalgesia and test its efficacy in differentiating individuals with OA who exhibit signs of centrally-augmented pain.

This wide-ranging review of the relevant background literature will start by providing a brief overview of pain terminology as applied in this thesis. Pain outcome measures will then be reviewed in Section 2.3. Pain is a complex amalgam of many different factors, with a wide range of tools used to measure different aspects of pain and nociception. This brief review will focus on the most commonly-used tools for self-report of spontaneous pain as well as quantitative sensory testing (QST) methods which are widely used to investigate the mechanisms involved in normal and abnormal pain processing.

Chronic pain is associated with poor response to standard treatment and this may indicate changes in pain processing which have lead to centrally-driven pain augmentation. Section 2.4 will discuss the evidence for the presence of centrally-driven mechanisms across a range of neuropathic, musculoskeletal and ‘central sensitivity syndromes’ such as fibromyalgia. This will include a review of current understanding about the neurophysiological mechanisms involved in acute nociception and the changes that drive peripheral sensitisation, central sensitisation and the more entrenched phenomenon of central augmentation. Osteoarthritis is a musculoskeletal condition in which signs of central sensitisation and central pain augmentation have been reported. Given its widespread incidence, OA is a useful pain model for the current investigation and so the evidence for altered pain processing in OA will be considered in detail.

Cold hyperalgesia has been proposed as a key indicator of altered central pain processing. Section 2.4 will review the evidence for this proposition. Basic science and psychophysical literature will then be reviewed in order to provide an understanding of the potential mechanisms by which cold sensation and cold hyperalgesia are normally signalled and how this may be modulated by inflammation or supra-spinal CNS changes. The assessment of cold hyperalgesia is problematic and so currently available testing methods will be considered and it will be proposed that more clinically applicable and reliable method for testing cold response is needed.
The final section (Section 2.6) will review the evidence for the use of topical menthol as an alternative method for assessing response to cold. Current psychophysical data using topical menthol as an experimental model will be reviewed within the context of the current investigation.
2.2 Pain Definitions

The terminology of pain research is not always universally agreed. For clarity, as far as possible, definitions will follow the current International Association for the Study of Pain Taxonomy (IASP, 2011 update).

Hyperalgesia versus Aldynia

Hyperalgesia is defined by the IASP taxonomy as “increased pain from a stimulus that normally provokes pain” (IASP, 2011 update). Therefore hyperalgesia is a more intense pain response than might be expected from a stimulus that has the potential to create pain. Hyperalgesia might present as an increased response to a normal or a supra-threshold stimulus. For example, if the temperature rises sufficiently high, heat will always create pain. Heat hyperalgesia therefore is a pain response to a temperature that is lower than would normally evoke pain.

The term ‘alodynia’ has created greater difficulties as far as the type of stimulus that may be considered to provoke an allodynic response is concerned. The term was originally coined to differentiate between an augmented response to a noxious stimulus that could potentially create tissue damage (hyperalgesia) and a completely abnormal pain response to a stimulus that could never create tissue damage. Alodynia is now defined by IASP as: “pain due to a stimulus that does not normally provoke pain” (IASP, 2011 update). To distinguish between hyperalgesia and alodynia, the IASP taxonomy advises that for alodynia the stimulus and the response should be in different modes, whereas for hyperalgesia they should be in the same mode. For stimuli such as light touch, this is unambiguous since by definition light touch is not capable of creating tissue damage. However some stimuli are difficult to define. For example, many authors use the term “cold alodynia” to mean a pain response to a cold temperature such as 20°C that is normally non-noxious and cannot create tissue damage. In particular, if the non-noxious cold evokes a paradoxical burning response it could well be argued that this is alodynic. However, IASP and other advice (Sandkuhler, 2009) is that: “(alodynia) should only be used when it is known that the test stimulus is not capable of activating nociceptors”. Alodynia therefore requires a loss of specificity of a sensory modality whereas hyperalgesia is an augmentation of the same sensory modality.

This current investigation will therefore use the term hyperalgesia to mean a pain response to any degree of mechanical, cold or heat stimulus. It is however an
interesting issue with regard to cold response as application specifically of the term “allodynia” to describe a paradoxical burning response to a non-noxious standardised cold stimulus could be a useful delineator.

Central sensitisation and central pain augmentation
There is also variable use of the term “central sensitisation” in current pain literature, which creates some confusion. Despite the IASP definition of central sensitisation as “increased responsiveness of nociceptive neurons in the central nervous system”, a number of authors use the word to include the additional complex of processes that occur at supra-spinal level and create an augmented pain state. IASP clarify that central sensitisation can involve both a drop in intrinsic nociceptor threshold to activation and increased post-synaptic responsiveness due to dysfunction of the endogenous pain control system (IASP, 2011 update). For the purposes of this investigation, ‘central sensitisation’ will be applied to mean the plastic sensitisation processes occurring at spinal cord level in the CNS that are self-perpetuating but reversible (Latremoliere & Woolf, 2009). The term “central pain augmentation” will be used to describe the more established state which is supra-spinally-driven where additional organisational or structural changes have taken place, and where allodynic responses are more likely to be observed (Latremoliere & Woolf, 2009).

Spontaneous pain versus evoked pain
In human studies, observable output (behavioural or communicated response) to a pain stimulus is a key component that informs about the mechanisms and processes involved. These pain stimuli may be associated with a clinical disorder in which case the term “spontaneous pain” is used. Improved control of clinical pain is clearly the ultimate goal in research and so understanding about the experience and impact of spontaneous pain is an essential component. The qualities of that spontaneous pain may also inform about mechanisms. However, spontaneous pain involves a large number of uncontrollable variables and so “evoked pain” is also investigated. This involves application of a controlled stimulus to either healthy or clinical populations and measuring the response to that stimulus.

Nociceptive versus neuropathic pain
There is much debate about these terms since “nociceptive” is defined as pain that derives from a nociceptive system that is injured (IASP), implying that nociceptive pain is synonymous with inflammatory pain. In contrast, “neuropathic pain” is defined as
pain that derives from a damaged or "abnormally functioning" (IASP) neural system. However, as will be expanded upon in the following sections, where augmented pain states are concerned, there may be some greying of these apparently clear definitions. It is clear that some individuals with conventionally-classified "nociceptive" pain disorders such as osteoarthritis (OA) report non-nociceptive qualities to their spontaneous pain (Hochman et al., 2011) and exhibit signs suggestive of an abnormally functioning neural system, including hyperalgesia to clearly non-noxious cold temperature. Disorders now classified as "central sensitivity syndromes" such as fibromyalgia (Clauw, 2009) may have elements of both nociceptive and neuropathic pain. For the purposes of this investigation, the term "neuropathic pain" will be applied to mean pain from identifiable structural damage to the nervous system. This may be identified clinically as deficits in sensory function such as reduced touch sensation. The term "neuropathic-type pain" will be used to describe the pain qualities such as burning or prickling that may be reported when no nerve damage is identified. The term "nociceptive pain" will be used when it is clear that the primary noxious stimulus is activating peripheral nociceptors in a neural system that is clearly undamaged.
2.3 Pain Measures

Introduction

Pain involves a series of complex inter-linked elements: production of nociceptive information (although not necessarily from tissue damage), transmission, processing, modification and interpretation of that information, followed by response output, ie the report of pain. Many of these elements may be measured although it must always be remembered that no element exists in isolation and must in some way be associated with the other elements for validity. This section will briefly explore some of the broader issues associated with human pain research. Then specific pain measurement approaches relevant to the following review of literature and later investigations will be reviewed.

Human pain studies may be perceived as scientifically problematic because they are reliant on information elicited from the individual and so lack objectivity. Self-report tools may ask a patient about the intensity or nature of their pain, psychophysical and brain imaging studies may provide a stimulus and then evaluate the individual’s response. All human pain behaviour is inevitably influenced by a complex array of social, cultural, and psychological factors. Testing methods that are designed to be more objective, such as QST must therefore be approached with equal caution because any test that seeks to assess a conscious human response must be influenced to some extent by behaviour-related variables. Pain research is therefore challenged by the need to control (or at least be aware of) these variables as far as possible. Equally it means that pain research involving participants and investigators from one social or cultural context may not readily translate to a different context.

Spontaneous versus evoked pain

Despite these provisos, the following sections of this literature review will show that there is a large body of data, both molecular and clinical, which provides broadly coherent, although sometimes divergent, evidence to explain how abnormal pain processing may develop and present in different clinical groups. Transmission, imaging and response studies tend to involve either evaluation of response to an evoked stimulus, or may involve spontaneous pain in clinical groups (symptoms or consequences of the every day pain caused by their pathology). There is considerable discussion about whether evoked and spontaneous pain are associated or whether they reflect entirely different phenomena. For example, if changes in the sensitivity of the pain system are related to overall pain experience, then assessment of hyperalgesia to
evoked stimuli should correlate with self-reported levels of pain. A number of studies have demonstrated a moderate, although not strong, correlation between mechanical hyperalgesia, measured as reduced pressure pain threshold, and every day pain intensity in individuals with knee OA (Arendt-Nielsen et al., 2010; Imamura et al., 2008). However, a brain imaging study by Parks et al. (2011) found that entirely different brain areas were activated by spontaneous pain and locally evoked pain at the knee in individuals with osteoarthritis. This paper concluded that, although evoked pain may inform about pain mechanisms, its direct relevance to the pain experience of the individual needs to be considered. Given this finding and the purpose of the current investigation in developing a new QST assessment, it will be important to ensure that there is a direct link between spontaneous pain report and test results.

**Gender influences on pain measurement**

Many studies have looked at whether there are gender differences in either pain processing or pain response. Despite the wealth of data there is still little consensus. Spontaneous (clinical) pain has been reported both as more prevalent and less prevalent in women than men. In a review of gender and pain, Fillingim et al. (2009) reported that studies which explore prevalence of neuropathic, musculoskeletal or osteoarthritic pain using self-report questionnaires, will report higher incidence in women. For example, in a multi-country study involving more than 85,000 participants, Tsang A et al. (2008) found that chronic pain was reported by 45% of women in contrast to 31% of men. Overall, it appears that women are more likely than men to report musculoskeletal pain, although there is limited evidence that this pain is of any greater intensity (Fillingim et al., 2009). There is also no indication in these studies as to whether these results reflect actual pain differences or gender differences in response adherence.

A number of studies have reported a significant difference between genders for mechanical and heat hyperalgesia. Chesterton et al. (2003) found that healthy males consistently exhibited 30% higher PPT at the first dorsal interosseous muscle than healthy females. This difference remained stable over multiple repeated measures. This finding was supported by the standardized reference values reported by Rolke et al. (2006). This study concluded that, since there was no gender difference for detection thresholds, distinct differences in mechanical and heat pain thresholds suggests that gender-mediated variations in central processing were likely to be the cause rather than differences in peripheral factors such as innervation density.
Few studies have compared gender differences in cold response. Wright et al. (2010)\(^1\) found no significant gender difference in contact thermode-tested cold pain threshold at a range of upper limb test sites. Other studies which have investigated cold sensitivity and gender have used the cold pressor (ice immersion) test. Fillingim et al. (2009) concluded that women tended to rate VAS pain higher than men and have lower tolerance (less time immersed) but there were no differences in threshold (time to first cold pain). Kim et al. (2004) reviewed the effects of gender, race, personality and various genetic markers on cold pressor tolerance and found that there is complex interaction between pain sensitivity and genetic variation, gender, ethnicity and psychological factors. Psychological variables therefore may mediate gender differences, although both anxiety and depression are phenomena more often reported by females (Hyde et al., 2008; McLean et al., 2011) so an association between these and pain is not necessarily causal.

It therefore appears that there may be gender differences in both spontaneous pain and evoked pain (and therefore in mechanisms contributing to the development of pain). However, it is extremely difficult to unravel the complex web of psychological, cultural and hormonal influences before even considering neuro-biochemical differences. On balance it seems wise to follow the advice of Fillingim et al (2009) and ensure that studies are as gender balanced as possible and that gender differences in key outcome measures are assessed or gender applied as a covariate.

**Tools for assessing pain processing**

A wide range of validated outcome measures have been used in studies that have explored pain processing mechanisms and pain response in humans. This section of the literature review will consider tools that have been used to measure the intensity and quality of spontaneous pain. Evoked pain measurement tools will also be reviewed.

**Measurement of spontaneous pain**

An individual’s report of their spontaneous pain its impact on their life is clearly an important outcome measure, since the majority of interventions will aim to reduce pain report and improve function and quality of life. Spontaneous pain is generally measured as an intensity at a particular location and as the quality of pain experienced. The body of literature relating to spontaneous pain measurement is vast. Since the

\(^1\) Poster included in Appendix 5
focus of this investigation is the development of a new test for evoked pain, the following review will only provide a very brief synopsis of some of the key factors.

- **Intensity of spontaneous pain**
  Spontaneous pain intensity is the most widely used measure for pain. These scales are attractive because they are simple to apply in many different research and clinical contexts. Visual Analogue Scale (VAS) or verbal analogue scales are by far the most widely used in clinical practice. The great advantage particularly of a verbal pain scale (0-10) is its simplicity and wide applicability. Bijur et al. (2003) found that a verbally administered Numerical rating Scale (NRS) 0-10 correlated strongly with a horizontal un-numbered VAS scale. Other studies have demonstrated that VAS rating correlates well with area of brain activity during experimental heat pain, suggesting that, in healthy subjects, VAS rating accurately reflects pain intensity (Valet et al., 2004).

However, there are a number of problems associated with these scales. A numerical scale that is intended to reflect a somato-sensory phenomenon such as pain will not be meaningful to all patients. Paice and Cohen (1997) assessed the use of a verbal rating scale to evaluate pain via telephone contact with cancer patients and found that, although the majority were able to use the verbal scale, those who could not rate their pain with this method had the highest intake of morphine pain relief. Williams et al (2000) interviewed 78 chronic pain patients who were using both VAS and NRS scales and found significant differences in understanding about the meaning of pain scale endpoints or interval markers between patients and between patients and clinicians. The authors conclude that a VAS rating is not so much a matter of matching an internal sensation to a number or distance, but rather reflects an individual’s private construction of meaning from a highly variable range of internal and external factors.

Kim and Buschmann (2006) reported that a faces pain scale was more meaningful for reporting pain intensity in older patients compared with conventional VAS scales. Gagliese and Katz (2003) also reported that VAS scales lack sensitivity when assessing pain post operatively. In particular the authors found that age differences in pain report were not differentiated with VAS intensity scales yet clear differences were found with a pain descriptor scale. In contrast, Bolognese (2003) found that VAS intensity and Likart-type pain scales were equally sensitive in determining change in pain following placebo or active drug intervention in older individuals with osteoarthritis (OA), both achieving an effect size of close to 1.
Hjermstad et al. (2011) recently reviewed the many pain scale studies and concluded that the many variations in scales and methodology made it difficult to confirm the validity or reliability of any one scale. Whilst many studies demonstrate that their pain scale variation is reliable between sessions, others suggest that different variations are not reliable. Inter-subject variability is undoubtedly an issue with VAS scales and so care must be taken when only a VAS pain intensity scale is used. Pain is not uni-dimensional and so it is preferable to use more than one scale when assessing pain (Hjermstad et al., 2011). Price et al. (1994) found that adding a scale for unpleasantness as well as for pain provided an improved degree of validity when measuring both clinical (spontaneous) and experimental pain.

Disease-specific pain scales such as the Western Ontario and McMaster Universities (WOMAC) indices aim to contextualise pain intensity ratings (Bellamy et al., 1988). The WOMAC scales have shown high levels of reliability and validity in measuring pain during functional activities that are specifically impacted by the joint involved and are considered to be a ‘gold standard’ self-report tool for OA. Other OA-specific pain and function scales have more recently been created. The OARSI-OMERACT initiative created an alternative tool for measuring OA pain that includes questions about the constancy or intermittent nature of pain. Hawker et al. (2008a) reported that the new tool had high reliability (r= .085) when administered over the phone and re-tested over a 40-96 hour time period. A recent study has also shown it to be responsive following pharmaceutical intervention (Bond et al., 2012).

However, in addition to measurement of intensity and constancy of everyday pain intensity in subjects with knee OA, Hawker et al. (2008b) also investigated the quality of pain and reported that, although the majority of individuals reported an aching, deep pain, there were some for whom pain had a more neuropathic-type quality. The authors concluded that additional information regarding pain quality would assist in fully understanding the pain experience for those with OA.

- **Quality of spontaneous pain**

Melzack (1975) was the first to develop a pain score that aimed to characterise the quality of pain. The descriptors section of the McGill Pain Questionnaire (MPQ) was developed following consultation with both patients with chronic pain and their clinicians. The 84 words listed are those that were more frequently chosen as
descriptors of the pain experience and are categorised according to sensory, affective and evaluative properties. Within these categories, each word is ranking according to the frequency with which it was chosen by the contributing patients and clinicians (Melzack, 1975). Sub-totals for each of the categories or a total Pain Rating Index can be calculated. The MPQ has been used consistently since its creation by Melzack and has shown high levels of validity and test-retest reliability. Grafton et al. (2005) evaluated the test-retest reliability of the MPQ in patients with OA and found high Intra-class Correlation Coefficients (ICC) for each sub-score, although the highest ICC being found for the sensory aspect of the score (ICC=.95). In contrast, current pain intensity showed very poor reliability (ICC=.75). This reinforces the value of including more than just a measure of pain intensity when assessing pain, particularly in a condition such as OA where intensity can vary dramatically from one day to the next (Allen et al., 2009). Strand et al. (2008) reported that the MPQ (Norwegian version) showed far better test-retest reliability with OA patients than with those diagnosed with other musculoskeletal pain conditions (not specified). This study found that the MPQ showed only reasonable responsiveness to change, although the study design did not appear to account for a number of extraneous variables, which may have influenced this result. The MPQ descriptor scales have also been used to characterise neuropathic pain (Mackey et al., 2012) although the MPQ is reported as having low discriminatory ability between individuals classified as having ‘definite’, ‘possible’ or ‘unlikely’ neuropathic pain (Rasmussen et al., 2009).

Although widely used for its proven reliability and validity, the MPQ descriptors list is disadvantaged by its extensive nature. The 84 words of the original list are challenging to use efficiently with patients and it has been proposed that a shorter list may be just as valid. Chanda et al. (2011) found that the MPQ was very reliable in discriminating between those with chronic low back pain (LBP) and those with sub-acute LBP, commenting that only a small list of words was needed to discriminate. The Pain Quality Assessment Scale (Victor et al., 2008) was developed to assess the type of pain experienced by individuals. In this scale each descriptor term is associated with a 0-10 numerical rating scale rather than just a tick box, meaning that pain experience can be characterised as composed of proportions of pain qualities rather just as all or nothing. More recently the PQAS has been modified to encompass sub-scales, in particular the sub-scales of ‘deep’, ‘paradoxical’ and ‘surface’ type sensations. The authors suggest that these sub-scales correlate with different pain disorders: for example, ‘deep’ sub-scale includes words such as throbbing and aching, which may be more associated with
inflammatory type pain, whereas ‘surface’ words are more associated with neuropathic-type pain qualities (Waterman et al., 2010). Lin et al. (2011) reviewed a number of pain descriptor scores and concluded that the PQAS was the most discriminative between different pain conditions, closely followed by the MPQ.

Other self-report instruments have been developed to apply pain descriptors as part of a diagnostic procedure. In particular there has been a recent surge in self-report scales which assess the presence of pain qualities commonly associated with neuropathic pain: e.g. PainDETECT and Doleur Neuropathique 4 (DN4) (Freynhagen et al., 2006; Bouhassira & Attal, 2011). These instruments do not assign rankings to words but ask an individual about the presence of specific neuropathic-pain type qualities. Scores above a cut-off indicate presence of neuropathic pain. The Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) weights the answers to five questions about presence of neuropathic symptoms with two physical tests of sensory loss or augmentation to provide a similar cut-off score to diagnose neuropathy (Bennett, 2001). Several of these neuropathic pain questionnaires have recently been used more widely to characterise pain in individuals with no diagnosis of neuropathy. PainDETECT in particular is becoming more widely used, partly due to its open access and its efficient administrative properties. PainDETECT is a single page questionnaire, which offers a simple way to calculate a total score and a suggested interpretation of whether the score implies ‘positive’, ‘negative’ or ‘unclear’ neuropathic diagnosis (Freynhagen et al., 2006). A range of studies have now used PainDETECT to assess the incidence of neuropathic-type pain in non-neuropathic conditions ((Steegers et al., 2008; Amris et al., 2010; Wylde et al., 2011; Hochman et al., 2012; Ohtori et al., 2012) and has opened up debate about whether presence of neuropathic-type pain qualities is necessarily diagnostic of nerve damage. It has been proposed that such pain qualities may also be reflective of centrally augmented pain (Gwilym et al., 2010). This is covered in more depth in section 2.4.

Measurement of evoked pain
Whereas spontaneous pain testing focuses on symptoms, evoked pain testing aims to assess mechanisms. The state of the pain system is evaluated by providing a specific stimulus, measuring response and comparing that response to normative values. Originally, these assessments were developed within the context of neuropathy in order to test the integrity of different neural pathways. Thus different stimuli test different pathways: mechanical, heat, light touch. Reduced response indicates damage
to the conduction pathway whereas an increased response, such as pain to light touch, indicates an abnormally augmented response. For example, chronic regional pain syndrome has been described in terms of characteristic combinations of sensory loss and sensory augmentation (Huge et al., 2008). It has been proposed that evoked pain sensitivity is associated with development of clinical pain states (Nielsen et al., 2009) and a number of studies have supported this by showing that elevated QST, particularly mechanical hyperalgesia, is associated with increased spontaneous pain (Arendt-Nielsen et al., 2010; Imamura et al., 2008).

A system of precise quantitative analysis of response to evoked stimuli has evolved: quantitative sensory testing (QST). This has been described as an extension of the normal neurological assessment procedure for somatosensory function (Rolke, 2006). Although a wide range of quantitative tests of response to a sensory stimulus have been developed, there have been calls to standardise both the tests, methodology and normative values, leading to the German Neuropathic network’s extensive work (Rolke et al., 2006; Maier et al., 2010; Geber et al., 2011). There now exists a widely accepted protocol for the QST tests that have been reported to be most reliable and discriminative (Rolke et al., 2006). In essence, QST involves applying a specific noxious or innocuous stimulus using a standardised protocol. This stimulus may be thermal, mechanical, chemical or electrical and is applied to a specific test site and a quantified response from the individual is sought. This might be to register that a particular sensation has been reached (eg pressure pain threshold) or it may be to provide a rating of intensity for a particular stimulus. Response is compared to normative values and classified as reduced or enhanced, thereby signalling loss of conduction or augmentation of signal. Choice of test site allows for further assumptions to be made: an elevated response at a site distant from the primary pain area may suggest central pain processing is involved (Suokas et al., 2012).

The method of limits is commonly used for assessment of sensory or pain thresholds and involves presentation of series of stimuli ascending / intensifying at a standardised rate (eg increasing temperature or pressure). Once the limit of threshold is reached, the stimulus returns to baseline. This form of testing has demonstrated reliability across a number of different studies (Moloney et al., 2012). However, there are disadvantages to this method of testing as it is highly reliant on the comprehension, concentration and reaction time of the individual being tested (Hansson et al., 2007). The method of levels is less frequently used due to the increased time required, but involves providing a
series of stimuli (within the same modality). The individual reports whether each stimulus is painful or not and the stimulus either increases or decreased accordingly until a limit of pain is reached (Hansson et al., 2007).

Geber et al. (2011) evaluated in patients with a variety of neuropathic diagnoses the test-retest and intra-observer reliability of the German Network’s bank of QST tests which involve a combination of method of limits and method of levels (thermal detection and pain thresholds, vibration threshold, mechanical detection (von Frey hairs) and pain (pin-prick) thresholds, and windo-up ratio (temporal pain summation)). This study found that all tests showed high reliability for both test-retest and intra-observer analysis (ICC = .80 to .93) at the painful site, although less good reliability at less painful or non-involved sites. Suokas et al. (2012) performed a systematic review of the efficacy of QST tests to discriminate between those with OA and matched controls. The authors concluded that, due to the poor reported methodology, only pressure pain threshold testing was adequately reliable and able to discriminate.

*Use of pain measurement during evoked pain*

Although QST using the method of limits usually only involves collection of a threshold value related to the stimulus modality, a few studies have suggested the additional collection of intensity or quality data to provide a more comprehensive picture of the sensory experience. For example, Quiton and Greenspan (2008) selected heat stimuli based on individual rating on VAS of either 70/100 for a high heat pain stimulus or 40/100 for a low pain stimulus. Test-retest results showed that VAS ratings to the same heat stimuli reduced over time, demonstrating a subtle habituation. Klepac et al. (1981) evaluated whether the MPQ was appropriate to use with experimental pain (cold pressor test) as well as spontaneous pain. The authors found that both intensity and quality of experimental pain correlated with MPQ Pain Rating Index, although not with Number of Words chosen, suggesting that changes in intensity are associated with lower ranked (less intense) word choice rather than complete change in quality and category of words chosen.

*Summary*

Pain involves the complex interaction between nociceptive processing, supra-spinal processing and interpretation of the nociceptive signal and final expression. A wide
range of reliable and validated tools currently exist that assess both spontaneous and evoked pain responses. Different pain measurement tools therefore may therefore evaluate pain mechanisms or may evaluate pain impact on an individual.
2.4 Abnormal pain processing in chronic pain
This investigation into a novel pain measurement strategy takes place within the context of the enormous impact that chronic pain disorders such as osteoarthritis (OA) has at both individual and societal levels. Chronic pain is a costly world-wide problem, burdening health and welfare systems and negatively impacting individuals’ quality of life for many years. Large-scale studies in US, Europe and UK have shown that approximately one in five adults suffers from chronic pain. A large telephone study of more than 46,000 participants from 16 European countries found that 19% adults reported on-going moderate to severe pain (>5/10 VAS) that had lasted more than 6 months (Breivik et al., 2006). OA was cited as the cause of this chronic pain by 34% of respondents, with spinal conditions cited by 15% and pain as the result of an injury by 12%. Individuals who experience chronic non-resolving pain require considerably more medical and welfare assistance, which ultimately carries a heavy financial burden. A recent health economics study estimated the annual cost of chronic pain in the US at $635 billion to the health system and a further $300 billion in lost productivity (Gaskin & Richard, 2012).

Pain as the disorder
Although the term ‘chronic pain’ tends to conjure up specific conditions such as low back pain, fibromyalgia or neuropathic pain, it can develop in a wide variety of neuropathic, vascular, visceral, or musculoskeletal conditions, which might be traumatic, degenerative or of surgical origin. Understanding about why some individuals develop non-resolving pain whilst others do not is still a matter of debate. However there is increasing acknowledgement in both the research and clinical communities that pain which persists beyond resolution of the original nociceptive problem is driven by changes in central processing (Phillips & Clauw, 2011; Woolf, 2011). For these individuals, pain has become self-perpetuating and is now the disorder rather than the symptom.

Pain that persists long after surgery is an example of this phenomenon. Although the majority of people recover well from surgery, severe persistent post-operative pain has been reported in up to 10% of patients (Haroutianian et al., 2013). It has been reported following a range of procedures including hernia, breast surgery, amputation, thoracic surgery and joint replacement surgery. Intensity of pre-operative pain is reported as a significant risk factor for chronic persistent post-operative pain and so it has been suggested that high pain barrage pre-operatively may already have triggered changes
in pain processing that is reinforced with the addition of acute surgical pain (Kehlet et al., 2006). Evidence of post-surgical neuropathic-type symptoms, for example with high PainDETECT scores, has led to proposals that peri-operative nerve damage adds to the abnormal nociceptive barrage which is then difficult to reverse (Steegers et al., 2008; Haroutiunian et al., 2013).

Examples from joint replacement surgery and amputations demonstrate that chronic pain post-operatively does not depend on nociceptive input. Up to 80% of patients who undergo limb amputation experience “phantom limb pain” and this appears to be strongly associated with intensity of pre-operative pain (Flor, 2002). Improved resolution of pain pre-operatively has been shown to reduce the incidence of “phantom limb pain” following amputation (Kehlet et al., 2006). Following total knee arthroplasty, where the painful joint is removed and replaced with a prosthesis, it is estimated that between 15% and 20% (Wylde et al., 2011) of patients continue to experience severe pain. Wylde et al. (2011) found that pain quality in those with persistent pain showed signs of altered processing: 13% of those with severe pain scored positive on PainDETECT and the most frequently selected sensory descriptors for pain quality were aching (45%), sharp (20%) and burning (20%). Associations between higher post-operative pain and widespread altered pain sensitivity have also been shown. (Dieppe et al., 2010) also used PainDETECT to assess presence of neuropathic-type pain at six months post knee joint replacement surgery and found 33% of those assessed had high scores associated with allodynia (no further details provided).

The development of persistent pain may therefore be associated with volume of afferent nociceptive input or with underlying abnormal pain processing, or a combination of the two. Before considering more specifically the evidence for abnormal pain processing in central pain syndromes and musculoskeletal disorders, the following section will explore the mechanisms involved in normal and abnormally augmented pain processing.

**The development of abnormal pain processing**

The experience of pain, whether acute and reversible or chronic and tenacious, involves a series of inter-linked elements: nociceptive signals from a peripheral source are received, converted into a signal and transmitted to supra-spinal levels. Here these signals are interpreted and modified (either enhanced or inhibited) in the light of multiple intrinsic and extrinsic factors. The individual's final report of the pain
experience is the end product, which is influenced by contextual, social, cultural and psychological factors.

The following section will provide background to this investigation by reviewing current knowledge about the neurophysiological mechanisms by which chronic pain is thought to develop, briefly considering normal nociception, peripheral sensitisation, central sensitisation and central augmentation. Although these stages overlap, for the sake of clarity the following review will describe the phenomena as if they were linear.

- **Nociception**

Understanding of the normal process whereby acute tissue damage is converted into a pain stimulus provides an important foundation for the understanding of the development of abnormal processes that drive persistent pain.

Acute pain is felt as an immediate, unpleasant sensation, often with a sharp, intense quality. However, this conscious awareness of pain involves a complex train of events occurring along the transmission and interpretative pathway.

Non-inflammatory nociceptive input causes activation of peripheral nociceptors at their axon terminals. Nociceptors are heterogeneous both neuro-anatomically and molecularly: Αδ-fibres are thinly myelinated afferents with higher thresholds and lower conduction speeds than non-nociceptive Αβ-fibres. Αδ-fibres can be sub-divided into Type 1 high-threshold mechanical nociceptors which normally respond to noxious mechanical input and high temperatures (>50°C), but which reduce their thresholds to both during inflammation (Basbaum et al., 2009). Type 1 fibres mediate the first sharp pain associated with pinprick-type input. Type II Αδ nociceptors have a very high threshold to mechanical pain but a lower heat threshold and so mediate first heat pain. C-fibres are smaller in diameter and unmyelinated, causing slower transmission speed and so signal poorly localised secondary ache sensation (Basbaum et al., 2009). Most c-fibres are activated by a range of stimuli. Sub-sets of c-fibres are also identifiable by their expression of different neurochemicals and cell membrane channels. For example, peptidergic c-fibres release inflammatory neuropeptides such as substance P and calcitonin-gene related peptide (CGRP). Voltage-gated sodium, potassium and calcium channels are also found differentially on c-fibres and each play a vital role in transmission of nociceptive signals from the periphery or in release of neurotransmitters to facilitate pain transmission (Basbaum et al., 2009). For example,
the sodium channel Nav1.8 is highly expressed by most c-fibres, as well as by some Aδ-fibres and has been implicated in cold allostynia since Nav1.8 antagonists are reported to attenuate cold allodynia (Attal et al., 2009).

Also expressed in Aδ and c-fibre cell membranes are receptor channels, particularly transient receptor potential (TRP) channels, which are responsible for transducing sensory and nociceptive stimuli. Originally thought to be functionally distinct and stimulus-specific, these channels are now known to be inter-dependent (Basbaum et al., 2009). TRPV1 for example was originally found to transduce noxious heat to provoke a sensation of burning pain. Subsequent research has found that TRPV1 is polymodal and can be activated by a range of chemicals including capsaicin, all resulting in the same burning heat sensation (Tominaga & Caterina, 2004). However, it is now known that TRVI is only one of many TRP heat receptor channels, each of which have been found to code specific temperature ranges (Dhaka et al., 2006). In addition, the co-expressed heat-sensitive potassium channel TREK-1 may modulate receptor activity (Woolf & Ma, 2007). Transduction of non-noxious and noxious mechanical hyperalgesia remains unclear and complex. Whilst low threshold Aβ-fibres signal non-noxious vibration and light touch, low threshold Aδ-fibres in hairs may detect light touch. High threshold Aδ- and c-fibre mechanoreceptors respond to noxious stimuli via a range of receptor channels, although the precise channels and mechanisms remain unclear (Campero et al., 2011; Reed-Geaghan & Maricich, 2011).

Aδ and c- terminate on interneurons in the more superficial laminae of the dorsal horn of the spinal cord, which are tonically controlled by a variety of local and descending inhibitory or facilitatory mechanisms. Consequently post-synaptic transmission of the nociceptive signal only occurs if either there is sufficient barrage or if central inhibitory mechanisms are themselves inhibited. Once a post-synaptic potential has been activated, transmission of nociceptive information occurs via the spinothalamic or reticulothalamic tracts (Basbaum et al., 2009). These terminate and interconnect with a variety of supra-spinal interpretation and control areas. There is no single brain area that is responsible for the sensation of pain (Wiech & Tracey, 2009). Rather pain is the result of the co-activation of multiple structures: those responsible for the sensory-discriminative aspects (such as the somato-sensory cortex), the amygdala and thalamus which process unpleasantness, the anterior cingulate gyrus and insular cortex which control emotional aspects. In addition, ascending information also connects with midbrain structures such as the periaqueductual grey, which controls the
balance of facilitatory or inhibitory influences at spinal level (Knudsen et al., 2011).

- **Peripheral sensitisation**
  Sensitisation of the peripheral nociceptor terminal occurs when inflammatory mediators and cytokines are released into the extra-cellular environment following tissue injury. This is experienced as hyperalgesia to local stimuli and a spontaneous pain quality that is likely to be localised and may be deep and aching (Basbaum et al., 2009). Peripheral sensitisation has been described as “stimulus-evoked plasticity” (Woolf & Ma, 2007) since it requires the maintained presence of nociceptive input. A wide variety of inflammatory mediators and cytokines are released from nociceptors or from non-neural cells such as macrophages and fibroblasts, which flood into the damaged area. These newly released signaling molecules, such as bradykinin, substance P, prostaglandins and cytokines such as Interleukin-1β (IL1β) or tumor necrosis factor (TNF), bind to specific cell membrane sites and enhance excitability of nociceptors (Dib-Hajj et al., 2009). Second-messenger systems within the cell terminal are also involved in this complex process, which in turn modulates the gating properties of ion channels and reduces the threshold to activation of receptor channels (Schaible et al., 2011). The increased sensitivity of multiple channels in the nociceptor membrane has a ‘snowball’ effect so that less nociceptive input is required for continued activation of the action potential. However, all of these peripheral processes are dependent on ongoing release of chemical mediators. As healing occurs, release slows and so channel thresholds gradually return to their usual level.

- **Central sensitisation**
  Early signs of central sensitisation develop alongside peripheral sensitisation, as a normal part of the healing process. However, sensitisation that does not resolve in parallel with tissue healing develops more firmly entrenched characteristics and is therefore less easy to reverse. Outward signs of central sensitisation include widespread hyperalgesia, temporal summation and increased size and intensity of receptive fields and enhanced persistent after-sensations (Woolf & Ma, 2007). The quality of pain may also change so that non-noxious cold may be felt as a burning sensation. It has been proposed that paradoxical burning is the pathological correlate of the “thermal grill illusion” first described by Craig and Bushnell (1994), whereby the normal inhibitory effect of Aδ cold-sensing fibres on heat, pinch, cold (HPC) c-fibres is itself inhibited in the dorsal horn. In this case, cool temperature that activates the HPC c-fibre is coded as painful heat (Nickel et al., 2012). Additional sensory and temporal
qualities to spontaneous pain have also been reported to change during central sensitisation, so that it can be difficult to differentiate with confidence between pain quality resulting from central sensitisation and from neuropathic damage (Amris et al., 2010).

Central sensitisation is the process by which enhanced pain processing spreads to the central nervous system, initially in the dorsal horn and then to supra-spinal pathways (Woolf & Ma, 2007). This results in enhanced pain sensitivity that spreads beyond the local area of injury. Central sensitisation may be activity-dependent in its earlier stages or may progress to an entirely self-perpetuating stage where no peripheral nociceptive input is required for pain transmission (Latremoliere & Woolf, 2009). The early stage is controlled by phosphorylation largely due to the effects of glutamate. The longer-lasting second phase is controlled by transcriptional changes which drive the synthesis of proteins required for ongoing phosphorylation (Latremoliere & Woolf, 2009). Central sensitisation is activated by an intense, sustained or repeated noxious stimulus. Intense noxious stimuli from muscles and joints have been reported to produce longer-lasting central sensitisation than from skin (Latremoliere & Woolf, 2009), thus explaining why central sensitisation is so prevalent following surgery or its association with joint problems such as arthritis. Glutamate is a key player in the process of central sensitisation. Activation of Aδ and c-fibres causes release of glutamate, substance P and CGRP at the central terminal in the dorsal horn. Glutamate binds to lower threshold AMPA receptors on central cell membranes. This normally initiate an inhibitory signal. However, during intense barrage, the magnesium ion which blocks activation of the normally silent high threshold NMDA receptor is forced out, enabling glutamate to activate the NMDA receptor and AMPA to become silent (Costigan & Woolf, 2000). The subsequent influx of calcium via the NMDA receptor causes dramatic intra-cellular changes in signaling cascades via self-perpetuating phosphorylation, including activation of protein kinase C, the mitogen activated protein kinase (MAPK) pathway and nitric oxide. As a result, previously sub-threshold inputs now drive post-synaptic action potentials and the dorsal horn is sensitised. This manifests as temporal summation, where a succession of relatively low intensity inputs steadily increases dorsal horn output and subsequent sensation of pain (Costigan & Woolf, 2000).

A range of other substances add to the central sensitisation mix. Substance P release in the dorsal horn also causes sensitisation of spinothalamic and spino-PAG projection neurons. Brain derived neurotropic factor (BNDF) and CGRP also cause central changes
which contribute to the overall reduction in activation thresholds and increased expression of pro-nociceptive molecules which perpetuate intense and long-lasting nociceptive transmission (Latremoliere & Woolf, 2009). The consequence of this is to change ion channel kinetics and properties, increase the density of receptors, and increase activation of kinases pre- and post-synaptically. Heterosynaptic facilitation may also occur during central sensitisation. If dorsal horn neurons are sufficiently sensitised, non-noxious Aβ-fibre input can result in a nociceptive signal and normally silent wide dynamic range neurons may be activated. This would be observed as tactile alodynia and secondary hyperalgesia (Costigan & Woolf).

Supra-spinally controlled changes in descending inhibition are also associated with this stage of developing chronicity. GABAergic and glycinergic inhibitory interneurons are deactivated, resulting in decreased tonic inhibition and therefore enhanced excitation of projection neurons, which then more easily transmit an ascending nociceptive signal (Basbaum et al., 2009). Inhibitory serotonergic, noradrenergic and dopaminergic pathways descend from the brainstem, in particular from the peri-aqueductal grey (PAG), locus coeruleus, the raphe nuclei, and the rostral ventral medulla (RVM) (Tracey, 2005). Following nerve injury or intense nociceptive barrage, the normal balance between inhibitory and facilitatory influence is changed so that the net result is an increase in facilitation from supra-spinal pathways (Basbaum et al., 2009).

*Centrally-augmented pain processing*

Central pain augmentation describes the point at which the CNS has made structural changes at both spinal and supra-spinal levels so that afferent input is no longer necessary for the perpetuation of the pain message. At this point the differences in spontaneous pain quality between neuropathic and non-neuropathic etiologies become particularly indistinct. Spontaneous symptoms may be intensified to include dysesthesias, shooting or electric-shock sensations and may be associated with additional widespread pain problems and alterations in mood, sleep or fatigue levels (Kindler et al., 2011).

A range of studies have shown that a number of morphological and functional changes occur in individuals with chronic pain. Although general tonic activation levels (assessed with cerebral blood flow) are higher in those with chronic pain, additional increases in acute pain brain areas are seen during episodes of increased spontaneous pain (Baliki et al., 2008). The sustained high level of back pain in this study was also
shown to increase medial prefrontal cortex activity, a brain area that involves the emotional, cognitive, and motivational processing of pain. This study also showed a strong correlation between activation of the medial prefrontal cortex and intensity and duration of the condition (Baliki et al., 2008). Volumetric studies using MRI have shown reductions in grey matter in individuals with chronic pain, generally finding a correlation between brain grey matter changes and the duration or intensity of pain (Baliki et al., 2011; Gracely & Ambrose, 2011). However, the cause of this apparent degeneration is unknown. Apkarian et al. (2011) have proposed that changes in grey matter volume are a sign of chronic nociceptive input-induced supraspinal neuronal degeneration. Imaging studies of migraine, fibromyalgia, OA and neuropathic pain patients have all demonstrated significant loss of grey matter in brain areas that are associated with the integration and modulation of pain signals. It has been suggested that tonic activation counteracting endogenous inhibitory processes, for example in the brain stem may deplete available molecules leading to reduction in synapses and gradual destruction of cerebral pathways (Baliki et al., 2011). However, several studies have shown that grey matter changes are reversible. Rodriguez-Raecke et al. (2009) found decreased grey matter in the anterior cingulate cortex, right insular cortex, dorsolateral prefrontal cortex, amygdala and brainstem in subjects with chronic hip OA pain compared with controls. Re-imaging of a subgroup of 10 patients who were completely pain free after total hip replacement surgery, showed grey matter increases in the ACC, DLPFC, amygdala, and brainstem at 6 weeks and 4 months post surgery. Gwilym et al. (2010) have shown similar findings, also in hip replacement patients. Other authors suggest that the decreases in grey matter shown with imaging could be due to a range of additional factors such as decrease in cell size, neural or glial cell apoptosis, or even changes in blood flow or interstitial fluid (May, 2008).

- **Additional supra-spinal influences on altered pain processing**

The perception of pain however is not necessarily linked in a linear manner to the noxious input, but is influenced by psychological variables. This appears to be particularly the case for those with demonstrated altered pain processing. For example, fear of pain has been shown to influence pain persistence. Crombez et al. (1999) showed that in individuals with chronic low back pain, pain-related fear is a better predictor of disability than pain intensity. High levels of negative affect and low positive affect have also been associated with persistent pain (Strand et al., 2007). Anticipatory anxiety, which may be innate or learned, may also increase the ‘nocebo’
effect and thereby increase activity in facilitatory descending mechanisms (Tracey, 2010).

Depression has been associated with a range of centrally sensitised disorders such as fibromyalgia (Klauengberg et al., 2008), IBS (Robinson et al., 2011) and persistent pain in OA (Wylde et al., 2011). Although the link between depression and higher levels of pain has not been clearly shown (Yunus, 2007), imaging evidence suggests that there is a clear link with increased activity in brain areas which may mediate the link. Schweinhardt et al. (2008) found that depressive symptoms were associated with a cluster of increased activity in the medial prefrontal cortex in people with rheumatoid arthritis and that this area of brain activation mediated the relationship between depression and number of tender points. Apkarian et al. (2011) have proposed that the association of pain with emotion and emotional memory in the limbic system is a key factor. fMRI studies have shown that increased limbic and prefrontal cortex activity is seen across almost all chronic pain disorders (Apkarian et al., 2011). At a functional cortical level, these changes influence increased descending pain facilitation via the brain stem so that a vicious cycle is created.
Conditions exhibiting signs of central pain augmentation

The neurophysiological evidence therefore suggests that pain augmentation may develop either from excessive ascending nociceptive input or from pre-existing dysfunctional supra-spinal processes, which sensitise and misinterpret even non-nociceptive information. In the ascending model, nociceptive barrage of high intensity or over long periods of time results in central sensitisation and eventual structural changes in pain processing that persist beyond resolution of the original pain stimulus. In the supra-spinal model, there is no identifiable provoking pathology, but evidence of pre-existing alterations in central brain processes. Costigan et al. (2009) describes this as "the autonomous amplification of pain".

Increasingly it seems likely that particularly resistant pain disorders may involve a combination of both ascending and supra-spinal influences. A central pain processing system that is already sensitised or lacks inhibitory control will not manage to control additional peripheral nociceptive input but instead will sensitise the system still further. This was shown by (Yarnitsky et al., 2008) in a study of persistent post thoracotomy pain where those pain-free patients who showed reduced endogenous pain mechanisms pre-operatively had the greatest likelihood of pain that persisted beyond normal post-operative recovery time. In clinical examples where both abnormal central processing and nociceptive barrage are identified, such as with knee OA or irritable bowel syndrome (IBS), it is impossible to assess whether altered processing was a precursor to the development of chronic pain or vice-versa. However, there is some evidence that a proportion of the pain-free population exhibit reduced pain inhibitory systems (DNIC) (Granot et al., 2008) and that this may be more marked in older healthy individuals (Riley et al., 2010). Our own studies (Wright et al., 2010a)² have shown that cold pain threshold is elevated in a proportion of healthy pain-free subjects with no history of chronic pain. Those with higher CPT described the quality of the evoked pain as burning and stinging. These findings suggest that a proportion of pain-free individuals exhibit alteration in their central pain processing. If future studies support this, it raises interesting, perhaps uncomfortable, questions about the pathogenesis of this predisposition: developmental, genetic, environmental or psychological.

For the current investigation, the focus will be on exploring the signs of augmented central pain processing across a range of clinical disorders. So-called 'central sensitivity

² Poster included in Appendix 5
syndromes’ describe perhaps the most extreme examples of central pain augmentation but are instructive for identifying similar characteristics in other pain conditions. The characteristics of neuropathic pain will then be reviewed since it is hotly debated as to whether presence of neuropathic pain always reflects neuropathic damage. Evidence for the presence of central augmentation signs in musculoskeletal conditions will then be considered, with an emphasis on osteoarthritis.

**Central sensitivity syndromes**

The archetypal example of a disorder where pain is driven by central pain processing factors and not by nociceptive barrage is fibromyalgia. Although initially dismissed as a non-medical somatoform disorder, fibromyalgia is now acknowledged as a complex condition characterised by multi-system problems including sleep and cognitive dysfunctions, fatigue, altered immune response and neuroendocrine and sympathetic nervous system problems (Clauw, 2009). Despite the presence of tender points, it has now been shown that ascending pain input does not drive fibromyalgia, but that augmented central pain processing results in widespread peripheral pain hypersensitivity. This in turn creates nociceptive afferent signals that perpetuate the central changes (Clauw, 2009). In addition to diffuse hyperalgesia and temporal summation (Staud et al., 2007), those with fibromyalgia have shown dysfunctional descending pain inhibitory mechanisms (Kosek et al., 1996; Gormsen et al., 2012) and biochemical and structural changes in the brain such as differences in opioid receptor binding and reduced brain volume in pain-related area (Gracely & Ambrose, 2011; Robinson et al., 2011). Increased sensitivity to cold has been reported in a number of fibromyalgia studies. Berglund et al. (2002) found that females with fibromyalgia had significantly elevated cold pain thresholds (mean 22°C) which they described as paradoxically as heat, in contrast to the matched controls who rarely used the word heat and had CPTs below 10°C. More recently Smith et al. (2008) found that fibromyalgia was associated with sensitisation to repeated cold stimuli, although the same was not found for repeated heat stimuli. Using the neuropathic pain PainDETECT questionnaire, Amris et al. (2010) found significant correlations between mechanical hyperalgesia, tender point count and PainDETECT score. Most frequently reported symptoms were thermal (including cold) pain, and spontaneous burning and pricking sensations. Those with fibromyalgia are also identifiable by the lack of pain efficacy of standard analgesics, including opioids but greater success with CNS neuromodulators such as pregabalin (Siler et al., 2011).
Woolf proposed that the concept of “fibromyalgia-ness” was useful to describe the continuum of chronic widespread pain disorders which appear driven by central mechanisms in a similar way to fibromyalgia. Accordingly, a range of other pain conditions, previously labeled ‘psychosomatic’ are now being recognised as pathophysiologically similar to fibromyalgia (Kindler et al., 2011; Phillips & Clauw, 2011). These “central sensitivity syndromes” include conditions as diverse as temporomandibular joint (TMJ) disorder, migraines, irritable bowel syndrome (IBS), interstitial cystitis and endometriosis (Bajaj, 2003; Yunus, 2007; Kindler et al., 2011). For example, IBS studies show low incidence of identifiable structural pathology (O'Connor et al., 2012) and signs of central sensitisation (Clauw, 2009). Other studies have shown local (inter-rectal) and widespread hyperalgesia strongest at the lumbosacral area but spreading to the cervical region, increased sensitivity to heat and cold (Yunus et al., 1989), inefficient DNIC mechanisms (Piche et al., 2010) associated with diffuse hyperalgesia, and alterations in brain chemistry (Chen JY et al., 2011). Given the absence of pathophysiological cause for pain in many of these conditions it is likely that abnormally augmented pain processing was a precursor to the subsequent pain disorder. The reported clustering of pain disorders (eg fibromyalgia with migraines and IBS) also implies an underlying central influence (Yunus, 2007).

**Neuropathic pain**

Neuropathic pain has conventionally been described as a completely separate entity, with its own unique pathogenesis, pathological processes and sensory responses. However, a growing body of data has drawn attention to the apparent cross-over in certain pathophysiological processes and pain qualities between individuals with neuropathic pain and those with centrally augmented pain. There is consequently disagreement over the correct interpretation of some neuropathic-pain type symptoms.

The IASP defines neuropathic pain as "pain initiated or caused by a primary lesion or dysfunction of the nervous system" (IASP, 2011 update). The controversy lies with the word ‘dysfunction’, which implies that structural damage is not needed for neuropathic pain. Recent studies have applied neuropathic pain questionnaires such as PainDETECT to clinical populations with no identified nerve damage and found significant percentages scoring as ‘positive neuropathic’ (Hochman et al., 2011). This has caused greater controversy, with one group interpreting this as a demonstration of just how
widespread nerve damage is, whilst the other argues that it illustrates that neuropathic pain qualities are a sign of centrally augmented pain, not necessarily neuropathy.

In defence of neuropathic pain as a distinct entity, the patho-etioloogy of neuropathic pain involves features which are not described for other conditions. Neuropathic pain results from direct damage to nerve tissue and a number of features are unique consequences of nerve damage, described as a maladaptive response to injury (Costigan et al., 2009). A number of spontaneous sensory effects are reported in neuropathic pain. For example, ectopic foci are experienced as spontaneous shooting or electric shock sensations and are generated by activation of nociceptive impulse from within the neural system rather than from outside, as is the case for inflammatory nociceptive input. Ectopic activity can be initiated by altered processes in the dorsal root ganglion (DRG), or from the site of the nerve injury or can be generated by previously non-nociceptive neighbouring neurons (Costigan et al., 2009). Nerve injury also initiates peripheral sprouting of the injured axon. However the regeneration impulse is excessive so that uninjured axons are also stimulated to sprout. This can lead new Aβ-fibre axons to terminate on lamina II of the dorsal horn, causing mechanical stimuli to be coded as nociceptive (Costigan et al., 2009). Spontaneous paraesthetic sensations such as a sensation of ants crawling on the skin therefore tends to be associated solely with nerve damage (Baron et al., 2010). In addition, some QST changes are seen as reflective of neuropathic dysfunction. ‘Negative’ sensory signs are deficits of sensation indicating that axon conduction has been affected. ‘Positive’ signs such as hyperalgesia, allodynia or dynamic mechanical allodynia are concurrently found (Baron et al., 2010).

However, not all individuals diagnosed with neuropathy develop neuropathic pain. In fact the percentages in some neuropathic conditions are remarkably similar to the percentage of those with musculoskeletal disorders found to show neuropathic-type pain. For example, only 20% of those with diabetic peripheral neuropathy report pain and only 50% of those with post herpetic neuralgia, a condition assumed to be synonymous with pain (Bennett, 2006). It may be that there are considerable similarities in effect, if not aetiology, of pain between neuropathic and non-neuropathic diagnoses and the integrating factor may be development of central sensitisation. For example, temporal summation and reduced thresholds to heat and cold stimuli have been reported in both neuropathic and centrally augmented chronic pain conditions. For example, cold hyperalgesia has been detected in both peripheral and central
neuropathic pain. Baron et al. (2010) reported that 21% of patients with post-herpetic neuralgia exhibited cold hyperalgesia. Bilateral cold hyperalgesia has been reported in patients with unilateral carpal tunnel syndrome (de la Llave-Rincon et al., 2009) and also in about 20% of patients with central pain after a thalamic lesion (Greenspan et al., 1993). Reduced descending pain inhibition has not been widely reported in neuropathic pain conditions. Seifert and Maihöfner (2009) reported bilaterally decreased adaptation to painful electrical stimulation and significantly reduced DNIC effect in patients with lower limb chronic regional pain syndrome (CRPS). In contrast, Gormsen et al. (2012) used a cold pressor conditioning stimulus and found that subjects with neuropathic pain responded with 40% reductions in spontaneous pain, showing a normal endogenous inhibitory effect. In contrast, the spontaneous pain of fibromyalgia patients increased. This study suggests that for some individuals with neuropathic pain, dysesthetic and ectopic pain, allodynia and hyperalgesia are driven more by peripherally-driven or centrally-sensitised mechanisms rather than by changes to supra-spinal inhibitory mechanisms. This may vary between neuropathic conditions, explaining different study results.

Assessment of pain quality is often an integral component of neuropathic pain diagnosis. It has been proposed that neuropathic pain has a burning and/or shooting quality with dysesthetic sensations such as tingling, crawling, or electric shocks (Baron et al., 2010). Neuropathic pain questionnaires have demonstrated efficacy in differentiating between those with identifiable neuropathic damage and those with inflammatory-type pathology by their choice of a specific cluster of descriptors (Jensen et al., 2005; Victor et al., 2008). Indeed these questionnaires offer as validation the fact that these pain qualities were selected by more of those with neuropathic pain than those with ‘nociceptive’ pain.

However, contradictory evidence suggests that there may be considerable overlap in spontaneous pain quality between those with clearly diagnosed nerve damage and those without. Rasmussen et al. (2009) used the MPQ to compare the descriptor choices of subjects diagnosed as having definite, unclear and unlikely neuropathic pain by experienced neurologists using bedside clinical tests and experience. The only discriminatory word was throbbing, which was reported significantly more by those who were unlikely to have neuropathic pain. There was no significant group difference in choice of conventional neuropathic-type words such as burning, shooting or prickling. Other studies have shown that neuropathic pain questionnaires identify
significant proportions of individuals without physical diagnosis of neuropathic pain. For example, Koroschetz et al. (2011) used PainDETECT and found that 1434 fibromyalgia patients and 1623 patients with painful diabetic neuropathy selected very similar descriptors to characterise their spontaneous pain. Other studies have reported that a percentage of those with musculoskeletal pain disorders report clearly neuropathic symptoms. Hochman et al. (2011) found that 19% of a community cohort with knee OA scored in the positive neuropathic category for PainDETECT. Hawker et al. (2008b) reported that, although the majority of individuals with knee OA described their pain as aching or sharp, a proportion selected neuropathic-type words such as burning or radiating. Sterling and Pedler (2009) found that 34% of subjects with whiplash associated disorder scored as positive for neuropathic pain using S-LANSS. High score was related to more complex presentation with higher pain and disability and greater sensory hypersensitivity. Moss et al. (2010) reported that individuals with hip OA who exhibited widespread cold hyperalgesia (cold pain threshold >15°C) selected words such as burning and stinging to describe the sensation at CPT.

Studies which differentiate pain disorders according to aetiology and compare group means for pain quality may therefore be overlooking the significance of the similarities that exist between sub-groups. A review by Bouhassira and Attal (2011) emphasises that several studies have shown that the similarities in pain qualities between etiologically different conditions tend to outweigh the differences. This review suggested that differences in choice of pain descriptor, whether for spontaneous or evoked pain, is more likely to reflect differences in pain mechanisms than in aetiology. The authors propose that future studies assessing the efficacy of interventions would be more informative if subject groups were selected based on pain quality rather than aetiology.

In conclusion, although some aspects of neuropathic pain development are unique, the processes of centralised pain are the same for neuropathic and non-neuropathic disorders. The maladaptive processes associated with nerve damage are likely to provoke central changes more quickly than for inflammatory pain. Immediate and intense tonic nociceptive barrage followed by incomplete and aberrant healing creates intense abnormal signals from the periphery, which more quickly will result in biochemical and structural changes in the CNS. However, whilst phenomena such as spontaneous ectopic foci and dynamic mechanical allodynia do appear to differentiate neuropathic damage, pain quality does not.
Musculoskeletal conditions and altered pain processing

There is a growing body of evidence showing that signs of centrally sensitised and centrally-augmented pain are also observed in individuals with musculoskeletal conditions and is often associated with poor outcomes.

- **Signs of central sensitisation**

  Widespread mechanical hyperalgesia as a sign of central sensitisation has been reported for a range of musculoskeletal conditions. Hyperalgesia to mechanical and thermal stimuli at both inflamed and non-inflamed joints has been reported in rheumatoid arthritis (Hendiani et al., 2003; Meeus et al., 2012). A number of studies of unilateral lateral epicondylalgia have also consistently reported mechanical hyperalgesia that spreads bilaterally (Fernandez-Carnero et al., 2009a; Coombes BK et al., 2012). Fernandez-de-las-Penas et al. (2009) found that decreased PPT at both the C5-C6 zygapophyseal joint and the tibialis anterior muscle were associated with increased pain intensity and duration in patients with myofacial TMJ disorder. Decreased PPT at tender points in the shoulder region and also at tibialis anterior, has also been reported in shoulder impingement syndrome and was associated with increased reported shoulder pain (Hidalgo-Lozano et al., 2010). Chronic neck pain studies have shown local mechanical hyperalgesia in the trigeminal area (La Touche et al., 2010) and at the distant tibialis anterior site (Johnston et al., 2008). Mechanical hyperalgesia in whiplash associated disorder (WAD) has been widely studied. In a recent systematic review Stone et al. (2012) found 14 methodologically strong studies that reported decreased PPT at thoracic, upper limb or lower limb sites. Chronic low back pain studies which report PPT are less prevalent and also more contradictory. O’Neill et al. (2007) reported mechanical hyperalgesia at tibialis anterior in subjects with low back pain (LBP) compared with matched controls, yet there were no significant group differences in PPT at any spinal sites. A follow up study by Neziri et al. (2012) found that PPT closest to the most painful area of the low back was the best predictor of LBP presence, but more distant PPT was less predictive than temporal summation.

Thermal hyperalgesia appears to show less consistency between conditions or studies. Fernandez-Carnero et al. (2009b) reported no differences in heat or cold pain thresholds either locally or contralaterally in subjects with lateral epicondylalgia. Yet Coombes BK et al. (2012) reported that the severe elbow pain was significantly
associated with elevated CPT and HPT bilaterally in the same condition. Ruiz-Ruiz et al. (2011) also found that increased sensitivity in CPT and HPT was widespread and homogenous, in contrast to mechanical hyperalgesia which, although increased, varied between sites. Evidence for thermal hyperalgesia in LBP is also conflicting, although presence of elevated CPT is slightly more consistent than for HPT. Neziri et al. (2012) found that both HPT and CPT were poor predictors of the presence of low back pain. In contrast, Lewis et al. (2010) found significantly elevated CPT at lumbo-pelvic tender points and contralateral non-tender points in those with chronic low back pain as well as significantly increased CPT at a distant shoulder site. CPT was the only QST measure shown to differentiate presence of LBP, which the authors attribute to early signs of central sensitisation. Sterling et al. (2006) reported that CPT at one month following injury in subjects with WAD predicted pain and disability at two years, with an odds ratio of 1.1 to 1.13. A more recent systematic review by (Goldsmith et al., 2012) has supported this early evidence in WAD and a recent multi-centre study reinforced the efficacy of this prognostic model, with CPT predicting moderate or severe disability at 12 months post injury, with a receiver operating characteristic (ROC) curve of 0.89 (Sterling et al., 2012). Chien and Sterling (2010) reported that, whilst CPT was higher in idiopathic neck pain patients than controls, WAD patients showed the greatest elevation in CPT, in similar manner to the progressively sensitised PPT measurement also found in this study. The authors conclude that sensory hypersensitivity is therefore not an all or nothing condition but more a continuum of increasingly augmented pain processing. As sensory hyperalgesia increases, so do pain and dysfunction, and treatment success potentially declines. In one of the few studies to explore QST predictors of treatment outcome, Jull et al. (2007) found that increased CPT was also associated with poor response to physiotherapy treatment.

• **Signs of central pain augmentation**

Signs of more established central sensitisation and CNS reorganisation have also been reported in a limited number of musculoskeletal studies. Leffler et al. (2002) found that reduced DNIC mechanisms were associated with longer duration of rheumatoid arthritis. Increased severity of WAD has also been associated with reduced DNIC (Daenen et al., 2011b; Van Oosterwijk et al., 2012). In female subjects with TMJ disorder King et al. (2009) reported no change in heat pain during a cold pressor conditioning stimulus. There is little data available for spinal pain, although Neziri et al. (2011) reported no significant differences in CPM for either neck or low back pain patients, using a cold conditioning stimulus. A review of DNIC / CPM studies has
suggested that, although there does appear to be an association between reduced endogenous inhibitory mechanisms and chronic pain, there are problems with many study methodologies, so that results cannot be relied upon due to issues such as potential assessor bias or influence of extraneous variables.

It has been proposed that centrally augmented pain is associated with CNS changes in perception or body image in some musculoskeletal conditions, in similar manner to neuropathic conditions such as CRPS. However, this is by no means a universal finding. Linnman et al. (2009) found that WAD was associated with distinct areas of increased cerebral blood flow. Daenen et al. (2011a) reported altered body image in those with WAD but this was not supported by Pedler et al. (2013) who found that discrimination between right and left was not impaired. Wand et al. (2010) found no difference in tactile thresholds between those with chronic LBP and controls, but significant differences in two-point discrimination and graphaesthesia. Giesecke et al. (2004) reported decreased activity in the pain inhibitory portion of the PAG during noxious stimuli in those with chronic low back pain. Baliki et al. (2008) used fMRI to compare subjects with chronic low back pain, CRPS and OA and found that, although there were similarities in both grey matter loss and excessive activity in different brain areas, each condition had its own "morphological signature", perhaps reflecting individual physiological characteristics.

Few studies have looked at the quality of pain experienced by those with more severe musculoskeletal pain. Sterling and Pedler (2009) used the self-completed version of the Leeds Assessment of Neuropathic Signs and Symptoms (S-LANSS) and found that the 34% of subjects who scored positive for neuropathic pain also had the highest pain, disability, local mechanical hyperalgesia and cold hyperalgesia. PainDETECT has been used by several studies. Freynhagen et al. (2006) found that 37% of patients with LBP scored as "positive neuropathic". In a study of more than 1300 patients attending their family doctor or specialist for chronic pain of musculoskeletal aetiology, more than 1/3 reported sufficient neuropathic-type pain symptoms to be classified as "positive neuropathic" (Jespersen et al., 2010).

In conclusion, there is a large body of evidence for the presence of some element of central sensitisation and central pain augmentation in individuals with chronic musculoskeletal pain. The strongest evidence is from QST studies, where the presence of widespread and reversible mechanical hyperalgesia is regularly reported. This is
likely to be reflective of earlier stage central sensitisation. Widespread cold pain threshold appears to be present in the more severe presentations thereby reflecting more established central changes. Altered pain modulation is less clear and shows significant variability between conditions.

**Osteoarthritis and altered pain processing**

Osteoarthritis is one of the most prevalent chronic musculoskeletal disorders, affecting more than 7% of the Australian population and predicted to be the fourth leading cause of disability by 2020 (AIHW, 2007). Prevalence increases with age, with more than 50% of people over 65 having radiological evidence of OA in at least one joint (Felson, 2005). However, there is a well-established mismatch between radiological prevalence and self-reported prevalence (presence of pain and disability), making accurate estimates difficult (Pereira D, 2011). This mismatch also suggests that OA pain may not always be perpetuated by nociceptive afferent input.

- **Pathogenesis and pain processing**

Understanding about the pathogenesis of OA is constantly evolving. OA was historically considered a primary disorder of cartilage yet it is now acknowledged that a variety of risk factors and pathological processes contribute to its aetiology. OA is currently described as a heterogeneous group of disorders of synovial joints, each variant having distinct pathological causes but resulting in similar morphological changes, biochemical features and clinical signs (Dieppe & Lohmander, 2005; Kidd, 2006). Malfunctioning of chondrocytes is a key component, leading to cartilage damage alongside hypertrophic changes in the synovium and capsule, and subchondral bone degradation. Gradual biomechanical changes create further stress on dysfunctional joint structures (Abramson & Attur, 2009). Radiographic diagnosis of OA consequently requires structural change: narrowing of the joint space, changes in subchondral bone and the presence of osteophytes (Finan et al., 2012).

The mechanisms of pain production in OA are equally unclear. Early animal studies showed that nerve fibres from healthy joints were insensitive to pressure or movement (Schaible & Grubb, 1993) and the avascular nature of cartilage strongly suggested that OA was a non-inflammatory condition. Yet improvements in imaging techniques have uncovered rich innervation in subchondral bone, synovium and capsule as well as the existence of synovial inflammation in OA (Sofat et al., 2011). Inflammatory cytokines and mediators such as prostaglandin E2 have been found in joint structures (Abramson
& Attur, 2009) and correlations have been shown between self-reported pain and radiological evidence of synovitis (Sofat et al., 2011) Consequently it is now acknowledged that OA pain may come from nociceptive activity in the presence of injury-induced inflammatory processes.

Nociceptive barrage from these sources will therefore provoke peripheral sensitisation in the short term, and if long-lasting or intense enough may progress to the more prolonged phenomenon of central sensitisation. Harden et al. (2013) have even suggested that OA is an ideal ‘model’ for central sensitisation since large joints with long-term chronic inflammatory or nociceptive input are ideal candidates to produce the level of afferent input required to activate NMDA receptors and trigger central sensitisation processes.

However, clinically, there is often a mismatch between the degree of pain and the extent of joint changes or biochemical inflammatory markers in people with OA. Ohtori et al. (2012) found a negative correlation between increased neuropathic-type pain report (PainDETECT score) and Kellgren-Lawrence signs of joint damage or presence of joint fluid. Dieppe et al. (1997) found good correlations between radiological severity of subchondral or cartilage damage but no correlation between these signs and clinical severity as measured by pain VAS. In contrast, the large cohort study by Duncan et al. (2007) calculated a composite OA pain score and found that there was a correlation between pain and joint damage. Valdes et al. (2012) found that pre-operative radiographic severity of knee OA was linked to low likelihood of persistent pain following joint replacement surgery, yet Dowsey et al. (2012) found the exact converse. The association between biological markers and pain severity is equally contradictory. Lee et al. (2011) found significant differences between OA patients and healthy controls in IL6, a pro-inflammatory cytokine associated with chronic inflammation, yet no group difference following QST pain testing. Brenner et al. (2004) reported no association between levels of inflammatory biomarkers such as PGE2, nitric oxide (NO) or COX2 in blood serum, synovial fluid or synovial membrane and either radiographic severity or clinical signs. It therefore appears that the relationship between physiological or structural measures and clinical measures are complex although not entirely unrelated. Dieppe and Lohmander (2005) suggests that radiographic evidence of OA may predispose an individual to the symptoms of OA but that a complex range of additional factors drives pain severity.
• **Signs of central sensitisation**

Earlier studies, particularly those investigating pain, have regarded OA as a homogeneous condition and therefore used cohort means in their analyses. Indeed osteoarthritis is still described in many contexts as the archetypal nociceptive disorder, differentiation from which is diagnostic of neuropathic pain disorders (Jensen et al., 2005). Yet increasingly it is understood that OA is heterogeneous both in terms of pathogenesis and in terms of pain processing. There now exists a body of evidence to show that a proportion of subjects with OA exhibit signs of central sensitisation, associated strongly with increased pain, decreased function and lower quality of life.

Widespread hyperalgesia been widely reported in individuals with OA (Suokas et al., 2012). Bradley et al. (2004) demonstrated the sensitisation of mechanical thresholds at the shoulder and arm in people with knee osteoarthritis. Harden et al. (2013) found mechanical hyperalgesia at knee and elbow sites in those with knee OA. Arendt-Nielsen et al. (2010) reported significantly reduced PPT at multiple points around the peripatellar region which correlated with self-reported pain in those with knee OA, indicating a link between local pain and nociceptive sensitivity. This same study also found good correlations between pain and widespread reduced PPT at the forearm and lower leg, although no correlations with radiological findings. Reduced PPT, from the lumbar spine to the foot, was shown by Imamura et al. (2008) who reported good correlations between sensitised PPT and self reported pain (WOMAC) and lower quality of life (SF36).

Additional signs of central sensitisation have been reported in subjects with knee OA when compared with healthy controls. Lee et al. (2011) demonstrated sensitised PPT at multiple sites, including the thumb. VAS intensity response to heat stimulus at the forearm was also significantly higher in OA subjects, although DNIC effect during cold pressor conditioning stimulus was not significant. Our own studies have found widespread cold pain threshold to be elevated in those with OA compared with matched controls: at the forearm and heel in OA knee and the shoulder in OA hip (Moss et al., 2008; Wright et al., 2010b). However Harden et al. (2013) found CPT only significantly elevated at the affected knee and Wyld et al. (2010) reported no significant difference in CPT between controls and those with OA knee.

---

3 Posters included in Appendix 5
Enhanced thermal and mechanical wind-up in knee OA (Harden et al., 2013) and enhanced temporal summation of pressure pain (Arendt-Nielsen et al., 2010) has also been reported. Neogi et al. (2010) found temporal summation present in 40% of a very large multi-centre cohort (n=1633). An additional study by this group looked at whether this phenomenon was associated with dysfunctional motor activity around the knee and found that temporal summation was associated with an approximately 30% greater odds of increased knee flexor co-activation during knee extension (Neogi et al., 2011). Courtney et al. (2009) also reported centrally mediated motor changes with OA knee subjects showing an increased flexor withdrawal response compared with controls.

It could be argued that many of these ‘central sensitisation’ phenomena could equally be seen as signs of undiagnosed neuropathy. The structural damage caused by the OA process makes temporary or permanent nerve injury quite feasible. Neuropathic pain studies emphasise that both negative and positive signs would be anticipated if nerve damage has occurred, with sensory loss reflecting loss of conduction and positive signs such as spontaneous dysesthetic sensations reflecting maladaptive healing (Baron et al., 2010). A number of studies have reported some reductions in nerve function combined with enhanced function. Harden et al. (2013) reported hypoesthesia to cold detection, punctate detection and vibration sensations in combination with widespread mechanical hyperalgesia and enhanced thermal and mechanical wind-up. Wylde (2010) and Arendt-Nielsen (2010) also reported reduced cold detection thresholds alongside mechanical hyperalgesia and temporal summation. However, no OA study has reported wholesale conduction loss, so the question of whether the changes in nerve function shown in OA in fact reflects damage to A-fibres is still debatable.

- **Signs of central pain augmentation**

Although there is strong evidence that, for some individuals with OA, pain is perpetuated by spinally-mediated central sensitisation, there are also indications that changes in higher centre processing may also play a part in abnormally augmented pain. Failure of arthroplasty to resolve pain in at least 20% of knee OA patients has been mentioned. Gwilym (2008) has suggested that the phrase “phantom joint pain” is appropriate to describe the phenomenon of knee pain in the absence of an anatomical knee joint, implying that abnormal processing of afferent input at higher centres may drive pain in some individuals.
Abnormalities in descending endogenous inhibitory systems have been shown in people with OA in several studies. Kosek and Ordeberg (2000) demonstrated reduced DNIC effect in painful areas for patients awaiting hip replacement surgery compared with matched controls. Subsequent studies using similar methodology in patients with knee OA have confirmed this finding of reduced DNIC (or CPM) effect (Arendt-Nielsen et al., 2010), with the most recent study reporting that the conditioning stimulus increased pain sensitivity rather than reducing it in patients awaiting knee replacement (Graven-Nielsen et al. 2012). Quante et al. (2008) used magnetoencephalography and electroencephalography to demonstrate reduced endogenous pain inhibition in subjects with knee OA compared with controls. Kosek and Ordeberg (2000) found that dysfunctional DNIC normalised for people with hip OA following joint replacement surgery. Graven-Nielsen et al (2012) found similar normalisation of CPM in knee OA subjects 9-18 weeks following knee joint replacement. This was accompanied by PPT and thermal temporal summation that returned to a normal baseline also.

A number of studies have also now shown structural and biochemical changes in the brain in those with OA. Gwilym et al. (2010) found loss of thalamic grey matter volume in subjects with hip OA. In an earlier study Gwilym and colleagues (2009) also reported increased activation of the PAG in those with hip OA which correlated with reduced punctate pain thresholds and higher score on the neuropathic pain questionnaire, PainDETECT. Previous imaging studies have proposed that elevated activity in the brainstem, which controls the balance between descending inhibition or facilitation of pain, is a marker for central sensitisation (Zambreanu et al., 2005).

Choice of neuropathic pain-type descriptors for everyday pain has also been proposed as an indicator of altered pain processing in conditions with no diagnosis of neuropathy. Amris et al. (2012) found strong correlations between PainDETECT score, tender-point count and pressure pain thresholds in subjects with fibromyalgia and suggested that PainDETECT was therefore indicative of centrally-augmented pain. Hochman et al. (2011) applied PainDETECT to individuals with OA and found that 19% scored as "positive neuropathic". Applying a modified version of PainDETECT (mPD-Q) alongside QST measures, this same group found that those with higher mPD-Q score had a higher odds ratio of meeting the criteria for central sensitisation (mechanical hyperalgesia and/or enhanced temporal summation and/or allodynia) (Hochman et al., 2012). Phillips et al. (2012) also found a good correlation between PainDETECT score and a composite score for widespread pain (number of pain sites x mean VAS
pain) in OA patients awaiting joint replacement. Hawker et al. (2008b) analysed the symptom characteristics reported as most distressing by OA subjects (open questioning) and found that aching and sharp were the most frequently chosen but that burning and radiating were also often chosen.

It is finally important to note that changes in psychological functioning have been reported as accompanying increased pain severity and altered inhibitory mechanisms in a number of centrally-enhanced pain disorders (Kindler et al., 2011). Wylde et al. (2011) reported that depression was significantly associated with ongoing severity of pain after total joint surgery and a systematic review of psychological factors influencing recovery after total hip or knee replacement surgery found that the only clear factors that predicted persistent pain were catastrophising and pre-operative diagnosed depression (Vissers et al., 2012). Murphy et al. (2011) used a composite measure of “centrally-mediated” symptoms (fatigue, sleep efficiency, and depression) as a co-variable and found that this accounted for 10% of the variance in pain severity after age and radiographic severity, in women with knee OA. A large European study of individuals undergoing hip joint replacement (Judge et al., 2011) has reported that higher positive pre-operative expectations, such as reductions in stiffness, predict post-operative success. The converse is also likely to be true. For example, it has been reported that persistent OA pain may for some be associated with unfulfilled expectation of pain relief, or ‘nocebo’, which is mediated via changes in brain-stem control over pain inhibition (Tracey, 2010). Activation of limbic emotion-driven pathways during spontaneous pain exacerbations has been demonstrated across a range of chronic pain conditions, apparently supporting this idea. The supra-spinal and spinal pathways that establish altered pain processing may therefore be directly influenced by a number of intrinsic and extrinsic psychological factors. Changes in inhibitory influence at spinal level are the likely consequence, leading to structurally maintained pain augmentation.

In conclusion, OA has conventionally been seen as a homogeneous disorder where nociceptive barrage from inflamed and damaged structures create localised sensitisation, which is ameliorated in the short term by anti-inflammatory interventions and in the long-term by surgical intervention. However, the reality is somewhat more variable and a percentage of those who undergo expensive joint replacement find no pain relief years after surgery. A body of evidence now exists to show that for some people with OA, central sensitisation and centrally-augmented pain
processes dominate, meaning that pain is self-perpetuating and unresponsive to standard treatments.

**Summary**

In summary, the signs and symptoms of altered pain processing have been reported across a wide range of disorders of neuropathic, musculoskeletal or unknown aetiology and there appears to be a unifying phenomenon whereby pain becomes the disorder. The process by which this develops may vary between pathologies. For some individuals, changes in central inhibitory control may pre-exist, causing the normal intensity of nociceptive input from tissue damage be augmented. For others, the intensity or duration of nociceptive input following tissue insult may be the trigger, resulting in early and intense central sensitisation followed by augmentation of pain. Although for this group pain augmentation is initially activity-dependent, over time self-sustaining processes at spinal and supra-spinal level develop and may include disinhibition (or active facilitation) of signals in the dorsal horn, functional or structural changes that convert non-nociceptive input to a nociceptive signal and increases in supra-spinal areas that respond to pain. In this way, centrally-augmented pain reflects development of a type of maladaptive ‘efficiency’ in nociceptive pathways. In some centrally sensitised syndromes it has been reported that additional sensory input such as sound or smell may also be enhanced (Kindler et al., 2011).

It is also clear that pain augmentation processes are rarely a matter of ‘all or nothing’. The development from easily reversible central sensitisation to entrenched central pain augmentation appears to be gradual and a number of studies have shown that it may be reversible. For example, Graven-Nielsen et al. (2012) have shown that altered descending inhibitory control and widespread mechanical hyperalgesia can normalise as early as nine weeks after knee joint replacement. Central pain augmentation therefore appears to be an identifiable continuum that may potentially be reversible in its earlier stages. Early identification of those showing signs of more established central sensitisation would therefore be desirable. In the case of individuals awaiting surgery, identification of those exhibiting central signs might result in postponement of surgery until pain had been reduced with pharmaceutical interventions.

Recent interest in identification of those exhibiting central sensitisation / augmentation has led to a number of different suggestions for clinical tests. Self-report questionnaires for neuropathic pain, such as PainDETECT or s-LANSS have been proposed as
assessments tools for centrally-augmented pain. Clinical test batteries have also been proposed. On research group and have designed and carried out initial testing on "bedside" test kits for assessment of neurological problems for those with post-herpetic neuralgia (Osgood et al., 2012) and for OA (Eaton et al., 2012). These include a series of self-report questionnaires and QST tests, the results of which are then used to classify individuals according to the number of 'normal' or 'abnormal' results. Nijs et al. (2010) have suggested that a combination of targeted history-taking, clinical examination that includes simple QST for sensitivity to heat, cold and touch at sites distant from the injury and analysis of response to treatment can be applied to diagnose central sensitisation. Smart et al. (2010) have similarly proposed a mechanisms-based clinical assessment approach to differentiation between predominantly nociceptive, peripheral neuropathic and centrally sensitised pain in individuals attending physiotherapy.

Clinical test batteries certainly show potential but mostly require considerable clinician skill in application or interpretation of the tests. Although a series of QST assessments would provide robust evidence, they are not clinically appropriate. The following section therefore considers whether a single component of QST might be selected to provide a valid but clinically applicable test of centrally-augmented pain. There is evidence to suggest that cold hyperalgesia may be a robust and key indicator of altered pain processing. However current techniques for assessment of cold response are problematic. The final section of this literature review considers whether sensory evaluation of the effects of application of topical menthol might be an effective alternative.
2.5 Cold hyperalgesia as an indicator of centrally-augmented pain

It has been proposed that widespread cold hyperalgesia may be a particular sign that central sensitisation has become more entrenched at both spinal and supra-spinal levels and so may indicate the early stages of centrally augmented pain.

The incidence of cold hyperalgesia

The presence of cold hyperalgesia at non-painful sites has been shown for a range of chronic pain conditions where signs of centrally-augmented pain are also well-established. CPT data for fibromyalgia is particularly strong. Fibromyalgia is described as the archetypal disorder of altered pain processing (Woolf, 2011) and many studies have reported significantly elevated cold pain thresholds at non-painful sites when compared with matched controls. For example, Desmeules et al. (2003) reported widespread cold hyperalgesia in individuals with long-standing fibromyalgia (mean CPT 17°C) when compared with controls (mean CPT 10°C). Kosek et al. (1996) also found that fibromyalgia patients had markedly higher CPTs at all sites but heat hyperalgesia only at painful sites. In addition to elevated CPTs (mean 24°C), Berglund et al. (2002) reported that all fibromyalgia patients reported paradoxical sensations in response to cold stimuli, perhaps indicating that this change in quality reflected changes in supra-spinal interpretative functioning. Smith et al. (2008) found that females with fibromyalgia showed sensitisation to repeated cold stimuli, although not to heat stimuli, in contrast to controls who show habituation to both. Cold, but not mechanical hyperalgesia, has been particularly associated with greater pain severity and disability and with additional central sensitivity phenomena such as poor sleep, anxiety and depression in fibromyalgia (Hurtig et al., 2001), implying that widespread elevated CPT may indicate long-term established central changes.

Cold allodynia is often reported as a cardinal sign in animal models of neuropathic pain yet, although it is widely acknowledged (Baron et al., 2010) incidence of cold allodynia is less explicitly reported. In a large European cohort study Maier et al. (2010) investigated the evoked pain responses of 1236 patients with neuropathic pain. Cold hyperalgesia was found in 22% of all patients with peripheral neuropathic pain, in 23% of those with central post-stroke pain and in 20% of those with pain after a thalamic lesion. Jorum et al. (2003) that subjects with post-traumatic neuralgia exhibited a mean CPT of 24°C. Chronic regional pain syndrome (CRPS) is partially diagnosed by presence of cold hyperalgesia (Vaneker et al., 2005; Baron et al., 2010). It is unclear why some individuals with neuropathic pain develop cold hyperalgesia, although central
sensitisation and spinal disinhibition are often proposed as key mechanisms. Central disinhibition has been suggested as a particular driver of cold pain in CRPS, in addition to changes in sympathetic nervous system function (Vaneker et al., 2005).

A number of studies have also found widespread cold hyperalgesia in subjects with persistently painful musculoskeletal conditions. Pain from carpal tunnel syndrome is often difficult to resolve and a number of studies have found a strong association between ongoing pain and bilaterally elevated CPT in individuals with unilateral symptoms (de la Llave-Rincon et al., 2009). Temporo-mandibular joint (TMJ) disorder has also attracted attention as it appears to involve central sensitisation driven by the trigeminal nucleus, somewhat like post-herpetic neuralgia. Pfau et al. (2009) found that those with TMJ disorder showed significantly elevated CPT at the trapezius muscle compared with controls. Park et al. (2010) found that pain origin made a difference: subjects with arthrogenic TMJ disorder were significantly more sensitive to cold pain at tibialis anterior (14°C compared with 4°C) as well as in facial areas (mean 18°C versus mean 5°C) whereas those with myogenic TMJ disorder had very similar CPT to controls at all sites. The authors proposed that chronic joint inflammation may produce more pro-inflammatory cytokines than myogenic pain so causing greater intensity of central sensitisation. Chronic pain from lateral epicondylalgia has also been widely associated with elevated CPT. An early study by Smith et al. (1999) found that widespread elevated cold pain thresholds were preferentially improved with regional guanethidine block, whilst having no effect on PPT, suggesting the influence of the sympathetic nervous system, as for CRPS. Bissett et al. (2012) reported that bilaterally elevated CPT was associated with greater disability and poor quality of life in those with unilateral tennis elbow whereas unilateral mechanical hyperalgesia was not associated. Ruiz-Ruiz et al. (2011) also found consistently elevated CPT across all test sites, in contrast to PPT which was highly variable.

Data from whiplash associated disorder (WAD) studies has provided strong support for the importance of widespread cold hyperalgesia as a sign of abnormal pain processing. Elevated CPT has been strongly associated with chronicity, severity and poor treatment outcomes in WAD. Sterling et al. (2006) found that CPT tested one month after injury was the best physical predictor of moderate to severe pain and disability at two years post injury, with an odds ratio of 1.1 to 1.13. This has been supported by other studies (Banic et al., 2004). Patients with chronic WAD pain who exhibited both widespread mechanical and cold hyperalgesia reported minimal relief after a 10-week
physiotherapy intervention, whilst those with fewer signs of sensory augmentation reported considerably better relief (Jull et al., 2007). Elevated CPT in WAD has been associated with self-reported neuropathic-type pain qualities (Sterling & Pedler, 2009) and in the acute stage, with pain catastrophising, although this is influenced by gender (Rivest et al., 2010). Additional studies have found central pain system changes in WAD, including reduced endogenous pain inhibition (Daenen et al., 2011b; Van Oosterwijk et al., 2012) and altered cerebral flow (Linnman et al., 2009). As a consequence a systematic review has concluded that cold hyperalgesia may be a particularly useful assessment of central pain augmentation in WAD (Stone et al., 2012).

Clinical and psychophysical data therefore point to the presence of cold hyperalgesia in a range of chronic pain conditions. The following section will review current understanding of the mechanisms by which cold sensation and pathological cold pain is signalled.

Non-noxious cold transduction

Cooling from a skin temperature (32°C) baseline can be felt within as little as 1°C change, yet the precise mechanisms by which cold is sensed are still elusive. Understanding is based on the fusion of findings from molecular, animal and human studies, which do not always correspond. It is understood that a complex of receptor and ion channels are involved although the precise mechanisms of extra and intracellular modulation remain unclear.

TRPM8 is a polymodal ion channel that is voltage-gated and activated by cool temperatures and a variety of cold-mimetics including menthol (Peier et al., 2002). Although TRPM8 predominantly evokes a mildly cold sensation in both animal and human models (Knowlton et al., 2011) it has also been associated with noxious cold sensation in certain circumstances (McCoy et al., 2011). TRPM8 receptors have been found on free nerve endings for Aδ and c-fibres in the epidermis as well as in the mouth and in the pulp and dentine of teeth. It has also been found in a variety of visceral locations such as in the bladder, prostate and respiratory system, but its role in each of these is currently unclear (Knowlton et al., 2011).

Whereas changes in heat sensation are transduced via activation of a series of TRP channels, each of which codes for a particular temperature range, TRPM8 is the only
Chapter 2

Literature Review

currently acknowledged cool-signalling TRP channel. This therefore raises the question of how different cool temperatures are signalled. TRPM8 from DRG cells has a temperature threshold to activation of around 25°C under normal conditions saturating at approximately 10°C (Knowlton et al., 2011). However the current produced by a single low threshold cold receptor is small so that firing of an action potential depends on additional extrinsic factors such as channel density and neuronal excitability. Excitability of the TRPM8 channels in each neuron terminal is controlled by the complex interaction of sodium and potassium channels with TRPM8 to either permit or restrict activation in a graded manner. It is this balancing act which probably controls how different cold temperatures coded (Belmonte et al., 2009). The precise molecular mechanism by which a cool temperature opens the TRPM8 channel is also unclear. Channel activation is dependent on the presence of the membrane lipid phosphatidylinositol-4,5-bisphosphate (PIP₂). Activation-induced influx of calcium ions through the channel causes a series of intra-cellular reactions, which culminate in PIP₂ being dislodged and the TRPM8 channel deactivating. This is the most likely mechanism for habituation to cold (McKemy, 2012) a phenomenon that has been shown to be deficient and even reversed in individuals with fibromyalgia (Smith et al., 2008).

Sodium and potassium channels also play important roles in the modulation of cold transduction. Influx of potassium (K⁺) into TRPM8-expressing neurons through any mechanism desensitises TRPM8 via intra-cellular mechanisms, thereby increasing the intensity of cold or menthol needed for activation (McCoy et al., 2011). In addition, K⁺ braking systems directly modulate the amount of cold needed to activate a cold-sensitive action potential. For example, co-located TREK-1 and TRAAK (K⁺) are opened by heat and the subsequent K⁺ ion influx gradually polarises TRPM8 requiring an increasingly intense cold temperature for activation. Cold on the other hand closes the TREK-1 and TRAAK channels, reducing K⁺ influx and enabling TRPM8 to be activated by a less intense temperature (McCoy et al., 2011).

One of the unusual characteristics of TRPM8 is that it appears to be able to transduce both noxious and innocuous cold. Both high and low threshold menthol and cold sensitive neurons have been found on Aδ and c-fibres but with expression characteristics which suggest separate roles for noxious and non-noxious cold transduction (Madrid et al., 2009). Low threshold cold neurons are more densely expressed in Aδ cold-thermal fibres, which are highly sensitive to menthol, implying
greater expression of TRPM8 channels, and express fewer of the K+ braking systems described above. These neurons are therefore likely to signal non-noxious cold (Belmonte et al., 2009).

**Noxious cold**

High threshold cold neurons appear to be candidates for transduction of noxious cold. They express fewer TRPM8 channels and have a higher density of K+ braking systems, meaning that a more intense cold or menthol stimulus is needed for activation. In addition, high threshold neurons express more Nav1.8, a tetrodotoxin-resistant (TTX-r) sodium channel which is largely expressed in peripheral neurons considered to be nociceptive (McKemy, 2012). Animal studies have shown that Nav1.8-null mice are unresponsive to noxious cold but show a normal response to non-noxious cold (Zimmermann et al., 2007). Whereas all other Na+ channels are inactivated during cold temperature, Nav1.8 remains open even during very intense cold. Cold even appears to reduce the activation threshold for Nav1.8, suggesting an essential cooperative role with TRPM8 in transducing intense noxious cold (Zimmerman et al., 2007; Abrahamsen et al., 2008). Given that Nav1.8 is expressed only on nociceptive fibres, it may be assumed that high threshold cold-sensitive neurons are c-fibres.

Very low cold temperatures will evoke a clearly painful sensation that may be slow to build, intense and aching (Beise et al., 1998). It is likely that in these circumstances cold pain is transmitted via deeply located polymodal nociceptors, activated by mechanical distortion and further sensitised by injury-associated mediators. Nav1.8 is also likely to be a key player due to its capacity to continue firing in extreme cold. However, healthy adults report low intensity painful cold at temperatures unlikely to cause tissue damage (Davis, 1998), with normal cold pain threshold reported at between 10°C and 5°C. Cold pain at these temperatures is still likely to be transmitted by peripheral activation of high threshold cold-responsive nociceptors (Campero et al., 2009; Serra et al., 2009). Psychophysically, painful cold is often experienced as a paradoxical icy and burning sensation (Chen et al., 1996; Morin & Bushnell, 1998), and so it has been proposed that the heat receptor TRV1 may be involved. TRV1 is exclusively co-expressed with pro-nociceptive substances such as substance P and NGF, this would therefore suggest involvement of nociceptors. Susser et al. (1998) compared conduction velocities with psychophysical findings using warm and cold stimuli and showed that the paradoxical heat sensation that accompanies noxious cold for some individuals is conveyed via the same slower c-fibres as heat and not by the faster Aδ-fibres. This also explains why paradoxical heat sensation tends to have a lag period
before being sensed (Susser et al., 1998). In another older but often quoted study, Craig and Bushnell (1994) proposed the existence of a population of HPC (heat, pinch, cold) nociceptors, which were responsible for the thermal grill illusion whereby alternating cool and warmth is felt as a burning sensation. This may well be the same population as found more recently by Campero (2009). Averbeck et al. (2012) recently reported an association between CPT and intensity of response to the thermal grill illusion.

Campero et al. (2009); (2010) used microneurography in humans to demonstrate the existence of two sub-sets of cold sensitive c-fibres. Whilst c-fibres were activated by cooling below 15°C, 'c2-fibres' were activated by low intensity cold as well as by menthol and warmth. Campero and colleagues (2009, 2010) proposed that c2-fibres play a complementary role to cold-activated Aδ fibres. In normal circumstances, non-noxious Aδ cold fibres inhibit c2-fibres. As the temperature drops the higher threshold c2-fibres are activated. The ratio of active c2 -fibres to Aδ fibres increases until eventually the C2 fibres are the only ones active. This idea is supported by earlier findings by Fruhstorfer (1984) and Susser et al. (1999) that A-fibre block changes a cool sensation to one of unpleasant icy, stinging and burning. Further cooling of the skin below 15°C triggers noxious-cold activated c-fibres and c2-fibres stop firing as a result of their activity-dependent slowing characteristics (Campero et al., 2009). This would provide an explanation for the experience of mildly noxious, unpleasant rather than painful sensation experienced by some healthy subjects at intermediate temperatures between 15-20°C. Campero describes this as a signal that the temperature is not yet noxious and dangerous but is starting to leave the “comfort zone” (Campero et al., 2009).

Pathological cold hyperalgesia

Cold hyperalgesia and cold allodynia, where the response to cold is either elevated or of abnormal quality, have often been reported as key symptoms following inflammatory and neuropathic damage, in both animal and human models (Caspani et al., 2009). Few studies have reported the exact nature of cold pain in these pathological conditions, but it appears that the quality of sensation is similar to that experienced during non-pathological cold pain (icy, burning or prickling/stinging). Once again however, the exact mechanisms are unclear but peripheral and central sensitisation processes, together with activation of the irritant-activated channel TRPA1, appear to be important players.
The processes involved in peripheral sensitisation have been described earlier in this review. For example, voltage gated Na+ channels become more active due to sensitisation by inflammatory mediators such as PGE2 (Schaible et al., 2011). Their expression is also increased due to the upregulation of proteins through the actions of peptides such as nerve growth factor (NGF) and substance P. This includes the cold-pain associated channel Nav1.8, which is also sensitised by NGF, PGE2 and TNFα (Schaible et al., 2011). Inflammatory-induced sensitisation of sodium channels in cold-sensitive Aδ, c2 and normal c-fibres will also occur so that a less intense cold stimulus will result in opening of cold channels and triggering of action potentials. To add further excitation, the effect of K+ brakes is also reduced during inflammation. However, the role of TRPM8 during inflammation is rather more ambiguous. Rather than being sensitised, TRPM8 is reportedly desensitised. In direct contrast to TRPV1, mediators such as bradykinin, and PGE2 initiate protein kinase C which then dephosphorylates TRPM8 (Babes et al., 2010). Yet in apparent contradiction, animal studies have reported that TRPM8-null mice show a clear reduction in response during inflammation, suggesting that the channel does play a role. The answer may lie in the mediation of the polymodal voltage and ligand-gated ion channel TRPA1.

There has been much debate over the role of TRPA1 in signalling cold sensation or cold pain in a non-pathological environment. However TRPA1 is emerging as an important signaller of noxious cold following injury and inflammation, being dubbed “the gatekeeper for chronic pain” (Bautista et al., 2012). TRPA1 is expressed in c-fibres alongside TRPV1 and is activated by heat and mechanical stimuli as well as by a wide range of chemical irritants including mustard oil, cinnamaldehyde, and allicin (Bautista et al., 2012). A great number of inflammatory substances also activate TRPA1, including bradykinin (Bandell et al., 2004). Second-messenger cascades involving G-protein coupled receptors modulate TRPA1 so that it becomes sensitised still further. However, the relationship between TRPA1 and cold remains elusive. Although it was originally reported that cold activated TRPA1, many studies have shown that blocking of TRPA1 by antagonist binding or use of knock-out mice, reduces the pain of inflammation but not the cold element (Chen et al., 2011; Knowlton et al., 2011). It now appears more likely that cold temperature modulates and sensitises TRPA1 but only once it has been activated by an inflammatory or environmental irritant (McKemy, 2012). In this way, cold pain is mediated by TRPA1. Once activated, TRPA1 also appears to play a crucial role in re-sensitising high threshold c-fibre expressed TRPM8, counteracting the
desensitising effects of bradykinin (Linte et al., 2007). TRPM8 therefore may contribute more of the intense cold sensation associated with cold hyperalgesia.

Much as neuropathic pain and cold allodynia are often associated clinically the exact relationship is less clear. Caspani et al. (2009) found that, although animals with chronic sciatic nerve constriction injury showed enhanced behavioural responses to cold, there was no concomitant increase in activity in TRPA1 or TRPM8 expressing neurons and in fact slight down-regulation of both channels in DRG cells. The authors concluded that the behavioural signs of cold hyperalgesia were caused by a central rather than a peripheral effect. Central sensitisation processes are therefore likely to play a significant role in the development of cold hyperalgesia.

It has been proposed that sensation quality reflects the transmission pathway (Hansen et al., 2007). Cold hyperalgesia, whether to non-pathological sustained cold in normals (Harrison & Davis, 1999) or a result of pathology (Berglund et al., 2002) has been consistently described as having qualities of icy, unpleasant or painful stinging or prickling. An increased stinging or prickling sensation associated with unpleasantness may suggest central augmentation of high threshold cold c2-fibres (Campero et al., 2009) as a result of NMDA phosphorylation during central sensitisation. Augmented cold signals from low-threshold Aδ-fibres may be felt as intensely cold or icy. Changes to the mode of sensation, such as with paradoxical heat or stinging in response to cold may also be an indication of long term heterosynaptic potentiation from established central sensitisation (Hansen et al., 2007). Alternatively, a number of different disinhibition mechanisms may be involved. Green et al. (2008) reported the presence of cold spots in the superficial skin that elicited a noxious stinging sensation when stimulated with non-noxious (26°C) cold that were easily inhibited by Aβ-fibre mediated pressure. The authors concluded that a similar response in those with cold allodynia may reflect dysfunction in central inhibitory mechanisms. Several studies have looked at the effects of ketamine and the μ-opioid alfentanil on neuropathic cold allodynia (Leung et al., 2001). Both studies found that both drugs had positive effects on cold allodynia, either decreasing CPT, pain intensity or allodynic area. Alfentanil is likely to exert its actions by reversing dysfunctional central inhibitory drive, either by increasing descending inhibitory drive from the PAG or by strengthening c-fibre inhibition in the dorsal horn (Leung et al., 2001) whereas ketamine reverses NMDA-mediated dorsal horn augmentation. Together these results suggest that cold hyperalgesia/allodynia is associated with both dysfunctional inhibitory drive and
central sensitisation in the dorsal horn. There do not appear to be any studies which have investigated DNIC / CPM integrity alongside altered response to cold but it might be hypothesised that the two may correspond. Tuveson et al. (2003) reported that cold pain threshold was the most stable sign of central sensitisation following hypertonic saline injection, leading the authors to propose that elevated CPT may reflect long-term central modulation, either by central pain augmentation or early changes to descending inhibitory controls.

In summary, there are recurring themes in the basic science, clinical and psychophysical literature, which suggest that cold hyperalgesia is particularly characteristic of an established centrally-augmented pain state. Widespread cold hyperalgesia is present across a range of poorly resolving conditions and has been linked to additional signs of established centrally-mediated pain in poorly resolving musculoskeletal and neuropathic conditions as well as in central sensitivity syndromes such as fibromyalgia. It has been proposed that cold hyperalgesia may signal established altered pain processing. The phenomenon is associated with central neuropathic pain, after stroke or thalamic damage and reductions in thalamic volume have been shown across a number of chronic pain disorders. Elevated CPT and changes in sensation quality have been associated with dysfunction in pain inhibitory systems, at dorsal horn and at supra-spinal level. CPT has been shown to improve after treatment with both the NMDA antagonist ketamine and also the μ-opioid alfentanil suggesting involvement of both central sensitisation and altered midbrain PAG function. It must be acknowledged that the current literature linking cold hyperalgesia with centrally-augmented pain is less than conclusive and more studies are required to investigate further this hypothesis.

Although widespread mechanical hyperalgesia is also a widely reported feature of centrally-sensitised pain disorders, it appears to be a less stable measure than cold and is more influenced by fluctuations in the peripheral and spinal neurochemical environment (Stone et al., 2012). PPT is more likely to vary between sites whereas there are rarely large temperature differences in CPT between sites, suggesting a stable and centrally-controlled phenomenon. Finally, cold pain assessment has the potential to provide more comprehensive information about the state of centralisation. Whilst heat pain evokes an almost universal quality, abnormally elevated cold pain may also be associated with the qualities of icy, stinging and burning in some individuals. It has
been suggested that identification of these qualities provides additional information particularly about extent of disinhibition and sensitisation.

**Assessing cold hyperalgesia**
Assessment of cold hyperalgesia would therefore be particularly valuable in a clinical environment for identifying centrally sensitised or augmented pain. However, there are difficulties in the currently available testing methods, both in terms of equipment and interpretation of results.

**Thermode-assessed cold hyperalgesia**
The majority of thermal testing is currently performed with a thermo-electric device, which creates temperature change using the peltier principle. When an electric current is passed across two dissimilar conductors either heat or cooling is produced, depending on the direction of the current. A peltier thermode cooled by water housed in metal casing that could be attached to the limb of a subject to assess sensibility to heat or cooling was first described by Fruhstorfer et al. (1976). The ability to reverse the current and probe temperature made thermal testing a less lengthy process and so the ‘Marstock method’ was developed. This is now also called the Method of Limits and involves presenting a subject with a temperature that usually changes at a rate of 1° or 2° per second (always in one direction for this method) and requiring them to register when the temperature has reached a desired threshold. At this point the current (and temperature gradient) is reversed. The desirable threshold may be when the subject feels first change in cooling or warmth (detection thresholds) or may be the point at which the cooling or heating becomes painful (pain thresholds). The quantitative output from threshold testing is a temperature, which is often recorded as the mean of several trials.

The technology of the peltier thermode means that heating is considerably easier to achieve than cooling. In particular cooling to a temperature that is universally noxious (below 5°C) is problematic as it requires an extremely efficient liquid cooling system and sophisticated choice of conductor materials to achieve a temperature close to 0°C. The two currently commercially available peltier thermodes use slightly different materials and technology and consequently achieve different minimum temperatures. The Somedic (AB, Sweden) thermode has a lower limit of 5°C. Studies that have used this thermode and include comparison with healthy controls report statistical
difficulties, as normal CPT for healthy individuals may be less than 5°C. Studies then may use 5°C as the lowest possible ‘normal’ value, which raises the normal mean and is likely to result in Type 1 errors when difference to the CPT of a clinical population is analysed. The alternative is to not analyse CPT results, and a number of studies do this, citing insufficient CPT data (Kosek and Ordeberg, 2000; Wylde et al., 2010). The more recently developed Medoc Thermosensory Analyser (TSA) (Medoc, Israel) is capable of dropping to 0°C. However, achievement of such as low temperature is still challenging even for a more efficient device and once the thermode temperature falls below 5°C the change rate slows significantly, particularly in a warmer environment, resulting in unintended sustained intense cold stimuli and potential variability in results.

Much as CPT testing is widely accepted as the optimal method for testing cold hyperalgesia, methodological issues exist, in particular relating to whether use of a single temperature value can be considered a valid reflection of cold hyperalgesia. The Method of Limits requires a subject to identify a sensory target (“cold starting to become painful”) from a series of descending stimuli. In addition to fully understanding the sensory target, the subject has to be able to concentrate on the constantly changing stimuli, be sensorily aware enough to identify rapid changes in sensation and decisive about when the target has been reached. This appears to be considerably less clear for cold than for heat pain threshold for most people with normal responses. It is likely that this lack of clarity reflects the more complex mechanisms by which cooling and cold pain are sensed. Increasing temperature activates a series of TRP receptors which ‘mark’ gradations of heat, and so is clearly felt. In contrast, cold requires integration of sodium and potassium channels, either one or two receptor channels controlled by intra and extra cellular influences and transmitted via a number of different Aδ and c-fibre pathways. This lack of clarity leads to high inter and intra-subject variability for CPT and so increases sample size requirements.

In summary, cold response is the most widely assessed in a research context using a contact thermode, but this method suffers from equipment and methodological problems. As a consequence cold hyperalgesia is often the least reported, or reported as the least reliable, QST measure. That being said, for identification of extremes the thermode is valuable. For example, an individual with a highly hyperalgesic response to cold (for example a CPT above 20°C) will be identified quite easily with the thermode, as will an individual who has a very low temperature threshold (below 5°C) who will
not respond at all. However, for more subtle discrimination between degrees of hyperalgesic response, the thermode leaves much to be desired.

The cold pressor test
The cold pressor test has also been used to assess cold pain thresholds and also cold pain tolerance. This method is less technologically sophisticated and involves submersion of a part of a limb (usually hand or foot) in an ice bath comprising ice and water, ideally maintained at a constant temperature, usually between 2°C and 5°C (Ruscheweyh et al., 2010). Threshold and tolerance to cold pain are measured as time, with pain threshold measured as the time at which the individual first signals that they feel cold pain. Although simple to administer, the cold pressor test is also problematic. The temperature provided by the ice bath is supra-threshold for most healthy individuals and so the test can only assess hyperalgesia rather than alldynia. The larger surface area in contact with this very cold stimulus causes increased temporal summation and so it is more an assessment of efficiency of pain mechanisms than of abnormal pain. It is widely reported that the cold pressor test is a potent activator of endogenous pain control and so ideal to use as the conditioning stimulus during DNIC / CPM assessment (Granot et al., 2008). A recent study by Ruscheweyh et al. (2010) compared cold response to the cold pressor test (at 2.5°C) with that to a sustained cold stimulus from a TSA contact thermode. To reduce variability, the hand was used for both tests and the thermode was set to oscillate around 3°C (to avoid habituation). Both tests were continued for 60 seconds and VAS pain ratings were taken every 15 seconds. The study found that pain ratings and consequent drop-out rate were considerably higher for the cold pressor test, resulting in a strong ceiling effect on mean values. Correlations between VAS values for both tests were highest in the early stages of the test, although overall correlation was still good (r=.70, p<.001). However the study found that the cold pressor test was associated with considerably more variance than the thermode test. In addition to factors mentioned above, the authors concluded that the high ceiling effect for VAS ratings is not ideal for differentiation of pain ratings and that the tolerance rating is influenced by a range of psychological variables. The authors concluded that, although the cold pressor test was more accessible (non-technical and inexpensive), its value lies mainly as a secondary assessment of pain sensitivity by testing efficacy of noxious inhibitory response.

Additional clinical test methods
Chapter 2

Two groups have recently developed 'bedside' QST kits that include a test for cold alldynia. (Scholz et al., 2009) reported on a clinical assessment tool composed of six questions and 10 physical QST tests which they named 'Standardised Evaluation of Pain' (StEP). The cold test involved the 10-second application of a brass rod kept at 20°C. Response measurement is unclear but hypoaesthesia to cold appeared to be defined as the subject being unable to feel any cold and an alldynic response as the patient reporting pain on contact. This study does not include data regarding reliability or discriminative ability of the cold element of the test so it is difficult to analyse efficacy. Baron et al. (2012) suggested that clinicians apply one of a range of cool stimuli intended to be at about 20°C (a glass with water in, a cool brass rod, or acetone) and ask a patient about the sensation they feel. Cold alldynia is diagnosed if the patient feels the sensation as painful. Eaton et al. (2012) and Osgood et al. (2012) briefly reported a "Bedside Sensory Kit" (BSK) that they proposed could be used to assess extent of neurological deficit (BSK-PHN) (Osgood) or neurological classification for patients with OA (BSK-OA) (Eaton). In both kits a brass rod (no other details provided) was used to assess cold alldynia. For the BSK-PHN pilot study (n=12), all tests, including the brass rod cold test, were reported to show ICC reliability >.70, but for the BSK-OA pilot study (n= 22), reliability was "inconclusive".

Each of these 'kits' attempts to produce an easy and inexpensive clinical tool but unfortunately their simplicity results in scientific imprecision, so that validity is difficult to discern. For example, none of the tests appear to include any standardised patient response, no test describes where to place the cool object, yet differentiating between local and widespread response is an essential factor in assessment of neuropathic or non-neuropathic pain augmentation. Whilst clinically applicable, it is difficult to imagine that these tests would be valid or reliable.

Maxwell and Sterling (2012) also described a simple clinical test for cold hyperalgesia but tested its efficacy against standard thermode assessment. In participants with WAD, a bag of ice was held against the neck by the investigator for 10 seconds and NRS rating measured. When subjects were grouped dichotomously according to CPT< or >13°C, receiver operating characteristic curve (ROC) analysis showed that NRS was significantly better than chance at predicting groups, although a specific NRS cut-off for cold hyperalgesia was less clear: a score <1 was likely to reflect a non-hyperalgesic response; a score ≥5 was likely to reflect cold hyperalgesia. This leaves considerable room for uncertainty so again may be useful in differentiating the most severe only.
Whilst acknowledging that the test was intended to be uncomplicated, there are problems with methodology that may influence the validity of the results. Use of ice to assess cold hyperalgesia is problematic because it is a supra-threshold cold stimulus for the majority of people. For those who are hyperalgesic it may become intolerable very quickly (as evidenced by those who had to drop out of the test before the 10 second application had finished). Application of ice during the cold pressor test has also been shown to trigger additional psychological responses that impact on pain assessment (Keogh & Mansoor, 2001). For example, few people are likely to be naive to ice and therefore positive or negative expectation of effect is likely to influence the pain rating result (Tracey, 2010). Physiological factors will have created additional variables. For example, cold transduction is influenced by changes in skin pressure (Green & Schoen, 2007) so holding the ice pack in position on the neck added potential interference by variable mechano-receptor activation. Presumed difference in application area size between the ice pack and thermode may also have influenced degree of pain summation between ice and thermode applications. None the less, this study by Maxwell and Sterling (2012) emphases the need for a valid and reliable cold pain test that can be applied simply and inexpensively in the clinic environment.

**A definition of cold hyperalgesia?**

The ultimate problem that QST assessment methods have is how to define hyperalgesia. Bearing in mind the issues with floor effects for healthy subjects during CPT testing and the ceiling effects for cold pressor testing, identification of hyperalgesia by statistical difference in mean from a control group is problematic. Based on individual study values, subjects who were considered cold hyperalgesic in one study may not have been so categorised in another study. For example, Moss et al. (2008) reports that the mean global CPT for individuals with OA was 12.9°C, which was significantly different to the control mean. Yet had these individuals participated in the study by Kosek et al. (1996) they would not have been categorised as cold hyperalgesic unless they had exhibited a CPT >15°C.

A more specific cut-off value might be more valuable for clinical purposes than a variable comparison of means. Sterling et al. (2011) have recently proposed a cut-off value of 13°C for cold hyperalgesia at the cervical spine in individuals with WAD, based on cluster analysis and ROC curves to predict disability. Although only group mean and 95% confidence interval data are available from other studies, it is clear that CPT values for widespread pain conditions such as fibromyalgia are considerably higher
than 13°C, often between 18°C and 24°C (Kosek et al., 1996; Desmeules et al., 2003). This difference is likely to reflect the extent of cold hyperalgesia in these populations but demonstrates the difficulty with a single temperature figure.

A recent cluster analysis assessment of data provided by different groups studying PPT, HPT and CPT in the conditions carpal tunnel syndrome, lateral epicondylalgia, TMD, whiplash and knee OA, found that 15°C was the best cut-off value across a number of disorders for differentiating according to severity of pain (VAS) or disability (using disorder-specific measures) (Wright, 2011). Figure 2.1 illustrates this.

![Figure 2.1: Percentage of subjects with CPT >15°C, using datasets from a variety of musculoskeletal disorders associated with chronic pain and poor resolution.](image)

From Wright, 2011 (with permission)

In addition to identifying those with greater severity of pain or disability, a valid cut-off temperature for cold hyperalgesia needs to clearly identify those whose response reflects abnormal pain processing. The cut-off temperature therefore needs to be one that is clearly non-noxious for the majority of healthy individuals but that is not so high as to only identify those with fixed central changes. The upper limit of normal CPT appears to be around 10-12°C (Bennett, 2006; Rolke et al., 2006). The lower limit of CPT for those with firmly established fibromyalgia or CRPS appears to be around 20°C (Bennett, 2006). A cut-off temperature in the range of 13-15°C therefore appears to be appropriate, with the more conservative 15°C being optimal.

The basic science and psychophysical literature also provides some support for 15°C as the most prudent temperature choice. Cold hyperalgesia represents the point at which a stimulus that normally would be felt as cool or cold, starts to develop unpleasant characteristics. Icy, stinging and burning qualities imply augmentation or misinterpretation of non-noxious cold signals, as the result of long term potentiation, central disinhibition or supra-spinal changes. Campero et al. (2009) reported that high threshold cold-sensing c2-fibres become active between 17°C and 15°C, beyond which additional polymodal c-fibres start to preferentially fire. In humans it has been estimated that this may normally occur once the temperature drops to below 12°C.
(McKemy, 2012), a value that would broadly fit with mean healthy CPT values. It would be reasonable therefore to predict that sensitised c-fibres might start to activate closer to 15°C, with c2-fibres also becoming active at a progressively higher temperature, perhaps 20°C and above. This also supports the use of 15°C as a temperature cut-off.

**Summary**

In summary, there is a body of clinical, psychophysical and basic science data that points to cold hyperalgesia as an important characteristic of pain that has developed centrally-augmented features. Current testing procedures have considerable flaws associated with equipment and methodology as well as inconsistent interpretation. This has resulted in cold hyperalgesia not being reported as frequently as mechanical or heat hyperalgesia. Although there are some recent developments in simple alternative methods for testing cold response, all are hampered by lack of precision in application and poor apparent reliability or validity. The final section of this literature review will therefore consider an alternative approach to assessing cold hyperalgesia.
2.6 Topical menthol as a test for cold hyperalgesia

A body of clinical and experimental evidence points to the potential importance of cold hyperalgesia as a measure of the centrally-driven processes that perpetuate chronic pain in many individuals. Pain centralisation has been shown to be reversible if identified and treated appropriately (Gwilym et al., 2010; Graven-Nielsen et al., 2012). A reliable and valid test for cold hyperalgesia might therefore allow identification of individuals at risk of developing centrally-augmented pain.

For research knowledge to be translated into the clinically valuable information, a test for cold hyperalgesia needs to be practically applicable in the clinical environment. The test needs to be inexpensive, so that all appropriate patients can be assessed. It also needs to be quick and easy to apply and require no complex equipment or skills. Finally the test needs to be straight-forward to interpret.

Current test methods

The currently available candidate QST tests are not ideal. Cold pain threshold testing, although the current “gold standard” and widely used, is associated with significant problems. The thermode equipment is expensive and there are problems reaching sufficiently low temperatures for interpretation of values against a normal range. As a consequence, CPT is often missing from statistical analysis and not reported because insufficient data has been produced by the equipment. Testing methodology is also problematic and there can be problems with variability (Goldsmith et al., 2012). Accurate and reliable CPT identification is challenging as it requires good understanding, sensory awareness and good reaction speed because the ‘target’ is essentially constantly moving.

Other approaches have sought to take a different approach whereby the subject is presented with a standardised thermal-cold stimulus and asked to record their response to it. These tests seek to overcome the technological issues associated with the thermode and provide a more clinically suitable application. The cold pressor test for example is easily accessible, requiring only a container, water, ice and a stopwatch (Ruscheweyh et al., 2013). However, there is consistent data indicating that cold pressor sensitivity is influenced by psychological factors more than other QST tests. In practical terms, the cold pressor test requires the entire hand or foot to be submerged so spatial summation becomes an issue and ceiling effects and early withdrawal are problematic. Indeed it is just this characteristic that makes the cold pressor test
particularly effective as a strong conditioning stimulus for DNIC/CPM testing. The further problem with any cold test that uses ice as the stimulus is that for many people, including healthy controls, the temperature of iced water (2° to 4°C) is already supra-threshold. As a result, discrimination between hyperalgesic and non-hyperalgesic individuals is more difficult because the stimulus is at a high level of intensity for everyone. A recently developed ice application test for cold hyperalgesia in WAD suffers from many of the same issues (Maxwell and Sterling, 2012). Withdrawal from the test or refusal to undertake the test could be seen (quite reasonably) as a positive sign of cold hyperalgesia. For those that are able to tolerate repetitions of 10 second ice application and provide an NRS value for the experience, the next problem is how to decide where the NRS cut-off between hyperalgesic and non-hyperalgesic should lie. Maxwell and Sterling (2012) applied ROC curves but found that the precise cut-off lay between 1 and 5/11. So once again the test showed poor sensitivity and was only able to discriminate between extremes.

Determination of a ‘hyperalgesic’ response is a problem common to all cold testing (and most QST) methodologies. Most studies currently solve this issue by comparison to a control group mean, which can vary by up to 10°C between different studies using different control groups, test sites and thermodes. The previous section considered in detail a potential hyperalgesic cut-off temperature for CPT. However, the psychophysical literature indicates that optimal identification of cold hyperalgesia requires a broader set of information than just a numerical value of temperature (CPT) or time (cold pressor). For good content validity, a test for cold hyperalgesia would need to include information about the type of sensation experienced during the cold stimulus: whether it was unpleasant or painful and whether icy, burning and prickling or stinging qualities were reported (Harrison and Davies, 1999).

An alternative approach to testing for cold hyperalgesia is therefore needed. The above summary suggests that a single stimulus test is the preferred option as it is simpler to apply and interpret in a clinical setting. However, to differentiate a hyperalgesic response, the intensity of stimulus needs to be carefully selected so that it evokes a non-noxious (cold) sensation for the majority of healthy subjects, rather than the supra-threshold stimulus that ice provides. However, if the stimulus is too mild, it runs the risk of only identifying the extremely abnormal response, such as with the cool brass rod suggested by Baron et al. (2012) or Eaton et al. (2012). Although possibly appropriate as part of a bank of tests for neuropathy, the cold hyperalgesia test of the
current study is designed to identify those in the process of developing centrally-augmented pain, with the goal of introducing more appropriate pain relief before symptoms become too entrenched. Consequently the test stimulus needs to be at a mild to moderate level of intensity in order to capture those in the earlier stages of centrally-developing pain as well as those in late stage.

**Menthol as an alternative test stimulus**

Cold temperature is not the only stimulus that can evoke a cold sensation. During the basic science investigations into the transduction of cold sensation, a large number of chemical agonists were discovered that activate the TRPM8 channel and evoke cold-type behavior in animals, including menthol, eucalyptol and icilin (Knowlton et al., 2011). Icilin has been used to promote cold behavior in animal models and is a potent agonist for both TRPM8 and TRPA1, requiring only very small quantities to activate TRPM8 in a calcium-dependent manner and also desensitising with prolonged stimulation in the same way as for cold. However, icilin is expensive and, crucially, has not been licensed for human use. A more practical option is menthol, the ingredient in peppermint that creates a cooling sensation. Very low concentration (<10%w/v) menthol is an ingredient found in items as diverse as foods, throat lozenges, shampoos, oral hygiene products, nasal decongestants, digestive remedies and cigarettes. More detailed pharmacological information about menthol can be found in Appendix 1.

Menthol appears to have distinctly different effects depending on whether it is used at very low or high concentrations. Low concentration menthol has been used for many years as a mild analgesic, for example in creams or sprays. Recently, potential mechanisms by which these analgesic effects are mediated have been reported by a number of studies. Animal studies suggest that analgesia appears to be mediated via opioid or non-opioid receptors in the superficial laminae of the dorsal horn, although there is debate as to which receptors are the key candidates (Galeotti et al., 2002; Proudfoot et al., 2006; Watt et al., 2008). There is also some disagreement about exactly what type of analgesia is provided by menthol and at what dosage (Klein et al., 2012), although this is complicated by data being provided from *in vivo* and *in vitro* rat and mouse as well as a few human studies. The finding that menthol is expressed differently in humans as opposed to mice or rats (Xiao et al., 2008) underlines the added difficulty in using animal models to explain human sensory findings. There does appear to be a general consensus on the efficacy of menthol in reducing heat or capsaicin-related pain (Green & McAuliffe, 2000), suggesting that menthol influences TRPV1
channels. The finding that dynamic mechanical contact reduces the cool sensation from 10% menthol (Green and Schoen, 2007) implies that the relationship with TRPV1 may be reciprocal and this is supported by Premkumar and Abooj (2013) who report a reciprocal relationship between TRPV1 and TRPM8 mediated by protein kinase C. It is unclear how human pathology may impact on this effect, although the mechanisms are clearly complex. Despite uncertainties about precise concentrations or mechanisms, current data suggests that menthol has a biphasic effect, and is able to effect both analgesic and hyperalgesic responses in humans and animals.

The psychophysical and basic science data is clear that menthol activates TRPM8 and in so doing evokes a cool sensation (Cliff & Green, 1994; Eccles, 1994). Green and Schoen (2007) showed that a 10% concentration of menthol applied topically evoked mild sensations of cooling and very mild noxious sensations, although no further details were given. Hatem et al. (2007) reported that 30% concentration evoked sensations of cool in 90% of subjects and warm in 10% following a 10-minute application. Basic science supports the notion that menthol evokes this cool sensation by activating TRPM8 cold channels. Peier et al. (2002) was one of the first groups to report that cold sensitive neurons expressing a particular cold-sensitive TRP channel were also activated by the cooling compound menthol. Bandell et al. (2006) showed that menthol binds to TRPM8 at a specific site that is quite separate to the cold gating site. Once bound, menthol modulates the ion conductance of the channel and increases voltage sensitivity. Janssens and Voets (2011) subsequently discovered that the channel is able to bind up to four menthol molecules and that with the addition of each molecule the open state of TRPM8 is stabilised. This may explain why in humans application of menthol has been shown to increase sensitivity to cold (Seifert & Maihofner, 2007).

The majority of psychophysical studies have applied high concentration menthol (40%) to both healthy and clinical groups in order to evaluate topical menthol as an experimental model of neuropathic pain. This equates to a noxious cold stimulus, akin to noxious (low temperature) cold. Seifert and Maihofner (2007) used the McGill descriptor index Pain Rating Index (PRI) score as a quantitative measure of sensation quality during their fMRI study of normal noxious and allodynic cold responses. They reported that PRI scores for noxious (individualised supra-threshold) cold were not significantly different to PRI scores for allodynic cold (induced by 40% topical menthol). This 40% noxious menthol stimulus is reported elsewhere to evoke a painful burning sensation, without altering skin temperature. (Wasner et al., 2004) found that
all subjects reported a cold-thermal sensation, starting about 2 minutes after
application and rising to a peak of 4.5/10 at around 9 minutes and staying at that level
until menthol removal at 20 minutes. In addition, 80% of healthy subjects reported a
painful burning sensation reaching a peak of 3/10 (NRS) after 8 minutes and
continuing at that intensity until menthol removal. No other sensations were evaluated.
Binder et al. (2011) reported slightly less intense findings with the same 40% menthol
application, with a burning sensation evoked at 5-minutes lasting until 20 minutes in
30% of all subjects. In addition, smaller percentages of subjects reported cold or
freezing sensations throughout the 20-minute application. However, there is limited
reporting of sensation quality in both studies.

Basic science does not provide much data to explain why high concentration menthol
evokes a burning sensation. Wasner et al. (2004) applied an A-fibre block before
applying 40% menthol and found that the cold sensation, together with mechanical and
touch sensation was absent in all but two subjects, although the burning sensation
intensified. This suggests that noxious response to high intensity menthol is mediated
by c-fibres, with the cold mediated by Aδ-fibres, as demonstrated using the same
method by Fruhstorfer (1984). Wasner does not report any icy or stinging or prickling
sensations but the study does not seem to have asked subjects for this level of detail
about the noxious quality.

So therefore, if the proposals of Campero et al. (2009) are also incorporated, a
theoretical model to explain the transition from predominantly cool and non-noxious at
lower concentrations to predominantly burning and noxious at high concentrations of
menthol may be proposed. Low concentration menthol is likely to activate
predominantly low threshold TRPM8-expressing cold-sensitive Aδ thermoreceptors
and so be coded as cool sensation. As the concentration of menthol increases, in a
similar way to gradual decrease in cold temperature, low threshold c2-fibres may be
increasingly activated so that the balance changes from Aδ to c2 activation. Given the
lower expression of TRPM8 and higher co-expression of TRPV1 and Nav1.8 in c2 cells
this stage is likely to be experienced as more intensely cold accompanied by the mild
but uncomfortable burning, stinging or prickling sensations that are described for
temperatures around 17-20°C as described by Campero et al. (2009). This stage is
described by Campero as a transition phase, warning that a change has occurred that
might become dangerous. As menthol concentration increases still further towards the
40% concentration described above, C2 cold-sensitive cells become less dominant,
although still active, as polymodal c-fibres, expressing TRPA1, TRPV1, Nav1.8 and additional pro-nociceptive neuropeptides become more active. The fact that noxious menthol (and noxious cold) is still experienced predominantly as an intense cold sensation indicates that TRPM8 continues to be activated. It is unclear whether Nav1.8 activation alone is able to signal cold.

It is likely that in the presence of inflammation or injury, the augmented effects of menthol will be mediated by similar mechanisms as for cold temperature. Circulating pro-inflammatory chemicals lower channel thresholds so that less menthol / cold is required for initiation of an action potential. It is understood that TRPA1 plays a significant role in signalling noxious sensation, particularly the burning quality associated with cold hyperalgesia during inflammation, as a result of its co-expression on c-fibres with TRPV1. A recent study has importantly found that human TRPA1 is activated by menthol (Xiao et al., 2008) so that, in the presence of inflammation, menthol is likely to evoke noxious burning sensations, in part as a result of TRPA1 activation. Although much is still unknown, several studies have suggested that menthol may have a greater potential for maintaining activation of TRPM8 and TRPA1 cells in the presence of inflammation. For example, change in pH inhibits the ability of cold temperature to activate TRPM8 (Andersson et al., 2004) whereas menthol is unaffected. Central sensitisation effects would be expected to influence menthol response in a similar manner to cold temperature response. Long term potentiation will augment all post-synaptic noxious signals, such as from TRPA1-expressing c-fibres or TRPM8-expressing Aδ or c-nociceptors. This would result in increased intensity of cold and noxious sensations to be coded in response to both cold and menthol. Finally, there is some evidence to suggest that menthol may be particularly influenced by the central disinhibition that accompanies central sensitisation and may be a sign of the development of more structural changes. In animal models menthol has been shown to be active at central spinal terminals (Proudfoot et al., 2006). As noted above, a number of different studies have reported that low concentration menthol is dependent for its analgesic effects on spinal pro-inhibitory channels in the superficial laminae of the dorsal horn, such as mGluRs (Proudfoot et al., 2006) k-opiods (Galeotti et al., 2002) and GABA-A receptors (Watt et al., 2008). During central pain augmentation, loss of these inhibitory substances is likely enhance menthol effects still further.
Summary

In summary, menthol activates TRPM8 receptors so that in normal circumstances a cool or cold sensation is evoked. In pathological conditions it is hypothesised that non-noxious concentrations of menthol will also signal cold hyperalgesia in similar manner to non-noxious cold. However, this hypothesis has yet to be tested since all studies applying menthol to clinical pain populations have used noxious 40% concentrations (Namer et al., 2008; Wasner et al., 2008).

However it is clear from the few relevant previous studies that once a suitable non-noxious concentration of menthol has been determined, an appropriate measurement method is essential to comprehensively characterise whether response is non-noxious or noxious. Psychophysical studies have shown that identification of cold hyperalgesia and of the mechanisms that may be causing that response requires both assessment of rating of noxious intensity and additional information about specific noxious qualities. In particular the qualities of iciness, burning, stinging and prickling are regularly cited as important distinguishing features. From a clinical perspective, the measurement system needs to be able to clearly identify an abnormal noxious response to a normally non-noxious stimulus.
2.7 Conclusions

In conclusion, there is a large body of evidence demonstrating that chronic pain is associated with signs of altered and augmented pain processing that is driven from central mechanisms. Centrally-augmented pain can develop from a wide range of neuropathic, musculoskeletal or surgical disorders and is associated with higher levels of self-reported pain and inability to function in daily life. Even in a disorder such as OA, which has conventionally been considered as the archetypal peripherally-driven nociceptive disorder, a significant percentage of individuals have been found to exhibit signs of central sensitisation and pain augmentation. Centrally-driven pain is a spectrum rather than an absolute, ranging from firmly-established, but still activity-dependent and reversible, central sensitisation to a disorder where there have been significant changes in brain structure and activity. At its most extreme, centrally-augmented pain disorders may involve a wide spectrum of clustered sensory sensitivities.

Identification of the degree of central sensitisation experienced by an individual would be invaluable in facilitating improved selection of the most appropriate intervention. However, currently testing is largely done in a research context and involves a wide range of QST measures and imaging, which is not practical in a clinical setting. Cold hyperalgesia has been proposed as an alternative stand-alone assessment that may identify centrally-augmented pain since the phenomenon appears to be centrally-mediated. Whereas peripheral and early-stage central sensitisation appear to influence mechanical and heat hyperalgesia, central changes to pain processing appear to be more influential for CPT.

Current cold response testing is limited, in terms of equipment reliability and methodological validity, but also in terms of clinical applicability. Menthol has been established as the primary agonist for the TRPM8 cold receptor channel and is associated consistently with cold sensations when applied topically. Menthol is a purely sensory stimulus, meaning that the problems associated with skin cooling are avoided.

This series of studies therefore aimed to investigate systematically whether a specific concentration of menthol could be used to differentiate normal and abnormal responses to cold in a valid and reliable manner.
Chapter 3

Study One

Evaluation of the liquid test formulation

3.1 Abstract

Background and Aims

Menthol is an established chemical agent that activates TRPM8, a channel receptor linked to cold perception. No previous studies have investigated in detail whether graded concentrations of topically-applied menthol evoke dose-dependent sensory effects in humans. Preliminary laboratory work showed that simple formulations release menthol across a semi-permeable membrane in a dose-dependent manner. This study aimed to assess whether three graded concentrations of menthol (low, medium and high concentrations) in solution would evoke similar dose-dependent sensory effects in healthy subjects. Associations between response to menthol and conventional thermal-cold pain threshold testing were also evaluated.

Method

A blinded, repeated measure design was used. Thirty-two healthy, pain-free subjects participated in the study. This sample size was based on previous topical menthol studies in healthy subjects: Hatem et al. (2006) used 39 healthy subjects but the study included insufficient data regarding response to menthol to enable a full power calculation to be made. Other studies included much smaller subject numbers but again without any power justifications: Wasner et al., (2004) 10 healthy subjects; Binder (2011) 12 healthy subjects. It was therefore concluded that a sample size of approximately 30 healthy subjects would provide a sufficient spread of responses in this initial study.

Each subject experienced each of the menthol concentrations plus the solvent control in randomised order over 4 test sessions. A standardised method of application and response assessment was used, developed in previous pilot work. Each concentration was applied to the volar forearm for 15 minutes. Assessments of intensity and quality of sensation were recorded every minute: three 100mm VAS scales were used to record intensity of cold, unpleasantness and pain; a McGill Pain Questionnaire descriptors list was used for subjects to select words to describe the sensation. Conventional cold pain
thresholds (CPT) were measured at the same forearm site using a peltier thermode for comparison with menthol evoked response.

**Results**
Significant concentration-dependent effects were seen for both intensity and quality of response. There were significant dose-dependent effects for cold, unpleasantness and pain VAS (p<.001 for cold and unpleasantness, p=.003 for pain) and time to onset for each sensation was significantly earlier for higher concentrations. There were also significant dose-dependent effects for descriptor indices: p<.001 for both PRI and NWC. More intense words such as icy, burning or stinging were selected at higher concentrations, with less intense words such as cool and tingling preferred at low concentration. Significant correlations were seen between cold pain threshold and sensation quality for low concentration Liquid A (r=.494, p=.004) and medium concentration Liquid B (r=.493, p=.004), but not for high concentration Liquid C (p=.879). When subjects were grouped dichotomously according to a CPT cut-off of 15°C, Liquid B showed the best discriminative ability: those with CPT>15°C reported significantly higher VAS for cold (p=.025) and unpleasantness (p=.050), significantly earlier onset of all VAS sensations, higher PRI score (p=.013) and were more likely to describe the menthol sensation as burning, stinging and/or prickling.

**Conclusion**
Topical menthol evokes significant dose-dependent sensory responses in healthy individuals, both in terms of intensity of cold and unpleasantness and in terms of the specific qualities of sensation experienced. Those participants classified as cold hyperalgesic using the conventional measure of CPT exhibited significantly higher levels of cold, unpleasantness and pain and higher PRI, suggesting a good association between measures. Practical difficulties with applying this low viscosity solution on a supinated forearm will need to be addressed in further laboratory studies.
Chapter 3

Evaluation of liquid formulation

3.2 Introduction & Background

It has been proposed that cold hyperalgesia may be an indicator of dysfunctional central pain processing (Berglund et al., 2002; Jorum et al., 2003; Stone et al., 2012). Conventional methods for assessing cold hyperalgesia are problematic due in part to equipment issues and in part to methodological factors (Chapter 2: 2.6). A test for cold hyperalgesia that combines simplicity with strong reliability and validity is therefore needed. This first study sought to investigate in vivo the initial validity of the proposal that topical menthol could be used to identify normal and abnormal responses to cold.

Current testing methods for cold hyperalgesia are not ideal. The cold pressor test is often cited as advantageous because it does not involve expensive equipment, requiring only ice, water, a container and a timing device (Ruscheweyh et al., 2010). However, tests involving ice are difficult to grade due to the inevitably supra-threshold cold temperature of the stimulus. The cold pressor test also involves immersion of the larger surface area of a hand or foot in cold, resulting in increased temporal summation and a high noxious effect, even in normals. The cold pressor test is therefore effective in providing a strong conditioning stimulus for assessing endogenous pain system efficiency (Arendt-Nielsen et al., 2008; Ruscheweyh et al., 2010) but is not so effective in clearly discriminating between hyperalgesic and non-hyperalgesic individuals.

The more widely used approach to assessment of cold hyperalgesia involves using a peltier thermode. This produces a steadily descending series of cold stimuli via an electric current passing across two dissimilar conductors (Fruhstorfer et al., 1976). Cold pain threshold (CPT) is the temperature at which an individual feels that the sensation starts to change from cold to painfully cold. Although the thermode allows for a more sophisticated and graded assessment of hyperalgesic response than the cold pressor test, there are difficulties in both equipment and method. Achievement of the low temperatures required for the full range of normal CPTs (<10°C) is technologically challenging and often not achieved. Consequently studies often report having to abandon CPT testing either due to equipment failure or due to the large ‘flooring’ effect, associated with the cut-off temperature of some devices, which results in compromise of statistical analysis. In addition, the requirement for a subject to be able to identify a specific target sensation (painful cold) during a constantly changing stimulus is challenging to achieve reliably. CPT standard deviations consequently can be large, both within and between subjects, once again often resulting in the dismissal of cold hyperalgesia as an unreliable test.
A fundamental issue with all existing cold response assessments is the lack of clear
definition of “cold hyperalgesia”. Studies which report cold hyperalgesia use the term
relatively, to describe whether the group mean for the healthy controls was
significantly different to the mean for the clinical group. Consequently the temperature
that would be defined as cold hyperalgesic in one study may be within normal range for
another study. Only one published study to date has attempted to identify a
temperature cut-off value for cold hyperalgesia. Sterling et al. (2011) used cluster
analysis and ROC curves to determine that an appropriate cut-off for cold hyperalgesia
in a whiplash population was 13°C. Additional work (Wright, 2011) applied the same
method to CPT data from a range of different chronically painful conditions and
determined that 15°C is a more appropriate cut-off value. Basic science would provide
theoretical support for a cut-off around 15°C (Chapter 2: 2.6), however there is as yet
no widespread agreement for cut-offs to use in psychophysical or clinical studies.

Aside from a specific cut-off value, the content validity of a single temperature value
representation of cold hyperalgesia is questionable. A normal response to cold appears
to involve changes to quality as well as the intensity of sensation that cannot be
determined from a simple threshold test. A number of psychophysical studies have
reported that a normal response to normally non-noxious cold temperature may be
described as dysesthetic (tingling, stinging or pricking) or as a paradoxical hot or
burning sensation (Davis, 1998; Harrison & Davis, 1999).

The current investigation therefore aimed to evaluate whether a fixed cold stimulus
from a specific concentration of topically applied menthol was able to identify
individuals with an abnormal response to cold. Menthol has been shown in vitro (Peier
et al., 2002) in animal studies (Andersson et al., 2004) and in human studies to elicit a
predominantly cold response (Green, 1992). A very small number of psychophysical
studies have used either low (Green & Schoen, 2007) or high (Wasner et al., 2004;
Binder et al., 2011) concentrations of menthol in the context of experimental pain in
normal subjects. These studies indicate that low concentrations of menthol evoke weak
sensations of non-noxious cool (Green & Schoen, 2007) whereas high concentrations
evoke noxious burning-cold sensations in a variable proportion of healthy subjects
(Wasner et al., 2004; Binder et al., 2011). In vivo therefore, high menthol concentrations
do not just increase the intensity of response but also appear to change the quality of
sensation experienced. Psychophysical studies therefore suggest that, as with cold
temperature, it should be possible to identify a menthol concentration that elicits non-
noxious cold sensation in the majority of healthy individuals but triggers a more
noxious cold response in a small percentage.

Basic science studies broadly support the proposal that menthol may evoke both non-
noxious cool and noxious cold sensations. TRPM8 thermal-cold receptors, expressed on
Aδ thermo-fibres, are also activated by menthol, evoking action potentials in a similar
manner to thermal cold stimulation and behavior similar to that produced by cold
stimulation in animals. Each TRPM8 channel has been found to include four menthol
binding sites, so that, as the concentration of circulating menthol increases, more sites
are filled, cell activation increases and the open state stabilises (Janssens & Voets,
2011). For TRPM8 channels that are expressed on thermal-cold Aδ-fibres, this is likely
to be experienced as increasing intensity of non-noxious cold. However, TRPM8
channels have also been reported as expressed on cold and menthol sensitive low-
threshold Aδ nociceptors (Campero et al., 2009). It is unclear under exactly what
circumstances nociceptive cold-sensitive neurons are activated but it indicates that
menthol (or cold) activation of TRPM8 may evoke more unpleasant sensations in
addition to cold. Noxious sensation accompanying cold may also be the result of TRPA1
activation. It has also been reported that menthol may activate human TRPA1 receptors
(Xiao et al., 2008), possibly only once the channel has been sensitised by other local or
environmental irritants, as occurs with cold temperature (Bautista et al., 2012). Since
TRPA1 channels are expressed exclusively on nociceptors (Kobayashi et al., 2005),
menthol activation of TRPA1 may evoke the more noxious sensations such as stinging
or burning.

In order to test more systematically whether menthol evoked a varied range of
intensities and sensations according to concentration, this study used three different
graded concentrations of menthol; low (Liquid A), medium (Liquid B) and high (Liquid
C). Preliminary laboratory work showed that simple formulations release menthol
across a semi-permeable membrane in a dose-dependent manner. The concentrations
were selected based on the sparse findings of the small number of studies that have
used menthol for experimental pain purposes (Green, 1992; Wasner et al., 2004; Hatem
et al., 2006; Green & Schoen, 2007; Binder et al., 2011). Although these studies provide
a general indication of effect, their focus was on the application of menthol to sensitise
the skin to other stimuli so that report of menthol effect is very non-specific. The
current study aimed to quantify specific menthol effects in a more systematic and
detailed manner in order to evaluate both effect and possible mechanisms. Response to
menthol in the current study was therefore measured using both intensity and quality measures. Intensity of cold, unpleasantness and pain were measured using standard VAS scales with simple end-point anchors marked maximum and minimum. This approach enabled a clear measure of whether each menthol concentration evoked non-noxious or noxious sensations and whether any individuals showed an elevated (hyperalgesic) response in terms of increased intensity. These three scales also allowed for some analysis of mechanisms: whether TRPM8 Aδ cold-thermal fibres only were activated or whether nociceptive processes had been engaged by each concentration.

Measurement of sensation quality would allow further analysis of mechanisms involved in the response of a healthy cohort to each concentration. The McGill Pain Questionnaire (MPQ) descriptors list has been widely used to characterise both spontaneous and evoked pain (Klepac et al., 1981). The list is composed of a series of systematically scaled descriptors, developed from studies with chronic pain patients and their clinicians. The MPQ has been shown to be valid and reliable in characterising and showing sensitivity to change in clinical pain states (Strand et al., 2008). In addition, the descriptors have been reported as reliable and valid for use with experimental pain (Klepac et al., 1981). The list includes both sensory and affective words which will enable the current study to evaluate whether there are systematic differences in response to each concentration. Word choice may also inform possible mechanisms. For example, affective words might indicate a more emotional supra-spinal influence over response; selection of purely cold-type words might imply predominant influence of TRPM8 channels with minimal c-fibre nociceptor involvement, since it has recently been shown that cold response is almost entirely mediated by TRPM8 alone (Knowlton et al., 2013); selection of dysaesthetic words might in contrast imply activation of nociceptors.

By definition, cold hyperalgesic response must involve a sensation of cold associated with some noxious sensation, both of which may be demonstrated by applying intensity and quality scores. However, the menthol test also needs to achieve criterion validity, by comparison with the current ‘gold standard’ test for cold hyperalgesia. Consequently the current study assessed CPT alongside menthol in order to examine whether there was an association between the two measures.

The primary aim of this first study was to evaluate the differences in intensity and quality of sensation between three graded menthol solutions applied topically for 15-
minutes to healthy subjects. The study also aimed to assess internal construct validity by examining the associations between the intensity and quality elements of the measurement. Finally, criterion validity was investigated by comparing response to each menthol concentration with conventional cold pain threshold value.

3.3 **Hypotheses**

1. There will be significant dose-dependent differences in intensity and quality of sensation experienced during the 15-minute application of liquid formulations containing different concentrations of menthol.

2. There will be significant dose-dependent differences in temporal characteristics between menthol concentrations: time to onset of first VAS rating, time to maximum rating and overall time-course of sensation intensity and quality.

3. There will be a significant association between measures of sensation intensity and measures of sensation quality at each concentration.

4. There will be a significant association between cold pain threshold temperature and sensory response to each menthol concentration.

5. Individuals with an elevated cold pain threshold (CPT>15°C) will exhibit an elevated response and different time to onset for each menthol concentration when compared to those with CPT<15°C.
3.4 Methods

Subjects
Thirty-two pain-free healthy adult subjects were voluntarily recruited for this study from the immediate surrounds of Curtin University, using posters and word of mouth. Inclusion and exclusion criteria were as in Table 3.1.

<table>
<thead>
<tr>
<th>Inclusion:</th>
<th>Aged 18-70(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Able to read and understand English</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion:</th>
<th>Current pain (\geq 1/10) VAS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Currently taking any analgesic or anti-inflammatory medication</td>
</tr>
<tr>
<td></td>
<td>History of systemic inflammatory conditions;</td>
</tr>
<tr>
<td></td>
<td>Neurological deficits (motor, cognitive or sensory);</td>
</tr>
<tr>
<td></td>
<td>History of other chronic pain disorders (e.g. fibromyalgia,</td>
</tr>
<tr>
<td></td>
<td>chronic low back pain);</td>
</tr>
<tr>
<td></td>
<td>Skin allergies;</td>
</tr>
<tr>
<td></td>
<td>Allergy to menthol or ethanol.</td>
</tr>
</tbody>
</table>

All participants provided written informed consent before participating in the study. Ethical approval was provided by Curtin University Human Research Ethics Committee (Approval number PT0058/2006).

Study design and procedures
This study used a within-subjects single-blind randomised repeated measures design to evaluate the sensory response of a cohort of healthy pain-free subjects to four concentrations of menthol in solution.

Subjects attended for testing on four occasions, each separated by at least 24 hours. Pilot testing indicated that carry-over of sensation from one menthol application interferes with additional applications even at different test sites in the short term. On each occasion a randomly allocated formulation was applied to the same volar forearm

\(^1\) The upper age limit of 70 reflects the need for the sample group to reflect the chronic pain population for whom this test may be appropriate. Previous studies have shown that in healthy individuals, age-related changes in pain threshold do not occur at this age (Lautenbacher et al., 2005; Wright et al., 2010).
test site for 15 minutes. Even though the ultimate goal of this investigation was to develop a test that could be applied during a short primary care (G.P.) appointment, a somewhat longer application time of 15 minutes was chosen for initial studies in order to allow sufficient time to explore the range of sensory effects that might be produced. In addition, this application time mirrored that used in previous studies (Wasner et al., 2004; Hatem et al., 2006), thereby enabling comparison to be made. On the first test occasion, before application of the first menthol formulation, cold detection and pain threshold testing was carried out using the same volar test site. All testing occurred in a laboratory with the ambient temperature maintained at 23-24°C. A volar site was preferred due to the lower variability between subjects in epidermal thickness and hairiness (Harrison and Davies, 1999). The dominant arm was used for all subjects.

Following a check of inclusion and exclusion criteria, subjects were asked some basic demographic questions. Subjects were then seated with their forearm positioned in full supination resting on a pillow placed on a table. A mid-volar 2x3cm area situated 5cms from the dominant arm wrist crease was marked. Before each test the site was gently cleaned with tepid water and dried with a paper towel. On the first test occasion, cold threshold testing was performed before any menthol testing since menthol has been shown to sensitise CPT (Wasner et al., 2004).

**Preparation of menthol solutions:**

All menthol formulation preparation was carried out under Pharmacy Laboratory conditions, following standardised procedures. In order to reduce variability between batches, a single 500ml batch of solvent was prepared and used for both the current liquid and gel formulation studies (Chapter 4). Four different solutions with graded concentrations of menthol were prepared, including a formulation with solvent only: liquid formulations A (low concentration menthol), B (medium concentration menthol) and C (high concentration menthol). Formulation D contained only solvent. Solutions were stored in opaque jars to prevent photo degradation and labeled to facilitate subject and assessor blinding.

**Outcome Measures**

All outcome measures were tested at the same volar forearm site. Cold pain threshold was tested at the start of the first test session. Cold detection threshold was tested before pain threshold to ensure intact neural pathways but results are not presented. Menthol formulations were then presented in randomised order, using a pre-planned
list created by dice throws, one formulation applied on each of the four test sessions. Subjects remained blind to concentration at each test session.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response to thermal cold</td>
<td>• Cold pain threshold (CPT) (°C)</td>
</tr>
</tbody>
</table>
| Response to menthol formulations A, B, C and D | • Intensity for cold, unpleasantness and pain (VAS /100)  
• Quality of sensation: Pain Rating Index (PRI) and Number of Words Chosen (NWC) |

1. **Cold Pain Threshold (CPT)** at the volar forearm was measured using a peltier thermode (Somedic, AB Sweden), a device commonly used and considered to be reliable and valid for the assessment of thermal QST (Rolke et al., 2006). The device uses the peltier principle to vary the temperature of the thermal plate by changing the intensity and direction of flow of current passing through (Shy et al., 2003). The thermal plate is housed in a probe, connected to the controlling computer, which is attached to the subject. The minimum temperature achievable by the Somedic thermode is 5°C. The 2x3cm contact probe was attached to the forearm test site with a velcro strap and subjects given several minutes to adapt to the baseline temperature of 32°C. The thermode temperature was set to drop at a constant rate of 1°C/sec. Standard Method of Limits (Fruhstorfer et al., 1976; Rolke et al., 2006), standardised instructions and standardised positioning were used. Cold detection threshold was always measured first in order to ensure intact sensory pathways before pain threshold testing. Subjects were instructed to depress the hand-held switch as soon as they perceived any cooling change from baseline. The temperature (°C) was recorded, with the thermode returning to 32°C. For cold pain threshold testing, subjects were instructed to press the control switch as soon as the cooling sensation changed to one of painful cold. This temperature was recorded (°C) and the thermode returned to baseline. One practice was followed by 3 trials, each trial separated by a randomly assigned pause of between 3 and 6 seconds, as for detection thresholds. The mean of the 3 trials was calculated for analysis.

2. **Menthol formulation tests.** Response to each of the menthol formulations (Liquids A, B, C and D) was assessed in randomised order on four separate test sessions. Subjects remained blind to concentration at all times. Once the test area on the forearm had been marked and gently cleaned with hypoallergenic soap and lukewarm water, a
2x3cm piece of cotton gauze was applied to the test site and 2mls of menthol solution dispersed across the gauze with a syringe. The test site was immediately occluded with a transparent adhesive dressing (Tegaderm, 3M) and the timer started for the 15-minute application. At one-minute intervals, subjects were asked to rate the intensity and quality of the sensation they were experiencing:

a. **Intensity** of cold, unpleasantness and pain experienced was measured by the subject marking each of three 100mm VAS scales on a single sheet of paper, one scale each designated for each sensation of cold, unpleasantness and pain. Each VAS scale comprised a 100mm line with end-points marked “maximum (sensation) imaginable” and “minimum (sensation) imaginable”. A new sheet with the three VAS scales was presented each minute.

b. **Quality** of sensation was measured by verbal selection of words (no minimum or maximum) from the McGill Pain Questionnaire descriptor list. The investigator recorded these choices on a data collection form at each minute. Subjects were also asked anecdotally if there were additional descriptors they would prefer to use and these were also recorded. MPQ Pain Rating Index and Number of Words Chosen index were calculated (Melzack, 1975). A higher score denotes a more severe response, based on data from 300 patients with chronic pain. This descriptor index has been reported as valid for both spontaneous and evoked pain (Klepac et al., 1981; Strand et al., 2008).

After 15 minutes, the dressing and gauze were removed, skin checked for erythema and the menthol solution immediately washed from the skin using hypoallergenic fragrance-free soap and tepid water, gently patted dry with paper towels. Further recording of VAS intensities and descriptors was taken at 20 and 25 minutes.

**Data Management and Analysis**

*CPT data:* A mean CPT temperature value was calculated from the three trials of forearm CPT. In order to evaluate whether those with thermal-cold hyperalgesia also exhibited a more intense response to the menthol test, subjects were then divided post-hoc into those with high and low CPT values. Two different methods were compared. Subjects were first grouped according to CPT < or > 15°C, as rationalised in Chapter 2: 2.6. This resulted in a high CPT group of 9 subjects (28.1%). Subjects were also grouped
according to dichotomous K-means cluster analysis, an approach previously used to
divide individuals according to CPT (Sterling et al., 2012). Cluster analysis resulted in
10 subjects (31%) classified in the high CPT group (cut-off value 14.3°C). Previous
studies would suggest that in a group of healthy subjects, a high CPT group of around
15-25% would be expected. A conservative cut-off of 15°C for CPT was therefore
applied.

*Menthol response data:* Data from intensity and quality ratings were managed as
follows.

a. *Intensity:* VAS values (0-100 mm) for intensity of cold, unpleasantness and pain
every minute for the 15-minute application were each quantified as:

- Area under the VAS-time curve (AUC) for 1-15 minutes, as a measure of
total intensity experienced during application;
- Maximum VAS value, as a measure of the greatest intensity experienced;
- Time to onset was calculated as the time at which the first VAS value
>0/100 was recorded;
- Time to peak VAS value was calculated as the time at which the maximum
VAS value was recorded.

b. *Quality:* MPQ descriptor list choices for quality of sensation experienced every
minute for 15 minutes were quantified using two standard MPQ indices:

- Pain Rating Index (PRI): sum of MPQ ranking values for each different word
selected. No single word was included more than once.
- Number of Words Chosen (NWC): the total number of different words
selected by an individual during the 15-minute application.

Normality testing was carried out using a combination of Shapiro-Wilk and visual
analysis of distribution. Cold VAS and PRI data were normally distributed (Shapiro-
Wilk p=.147 and p=.307) and parametric statistics applied. Unpleasantness and pain
VAS data were not normally distributed, so non-parametric analyses were applied.
When CPT data was divided according to CPT< or >15°C, data within each group was
normally distributed, allowing for parametric analysis of group difference. Data were
then analysed using the SPSS statistical package, version 19. Alpha was set at p<.05 and
the following analyses were carried out:
<table>
<thead>
<tr>
<th>Hypotheses and Related Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Difference in intensity and quality response between menthol concentrations:</td>
<td></td>
</tr>
<tr>
<td>• <em>Area under VAS-time curve for cold, unpleasantness and pain x3</em></td>
<td><em>Repeated Measures ANOVA</em></td>
</tr>
<tr>
<td>• Maximum VAS rating for cold, unpleasantness and pain x3</td>
<td>Or</td>
</tr>
<tr>
<td>• PRI and NWC quality ratings x3</td>
<td><em>Friedman ANOVA by ranks</em></td>
</tr>
<tr>
<td>• Specific word choice comparison</td>
<td><em>Descriptive comparisons</em></td>
</tr>
<tr>
<td>2. Difference in temporal characteristics between menthol concentrations: release times &amp; time-course:</td>
<td></td>
</tr>
<tr>
<td>• <em>Time to 1st sensation (mins) for cold, unpleasantness and pain x3</em></td>
<td><em>Repeated Measures ANOVA</em></td>
</tr>
<tr>
<td>• <em>Time to peak sensation (mins) for cold, unpleasantness and pain x3</em></td>
<td>Or</td>
</tr>
<tr>
<td>• <em>Time-course for each VAS sensation and key quality descriptors</em></td>
<td><em>Friedman ANOVA by ranks</em></td>
</tr>
<tr>
<td>3. Association between intensity and quality responses at each menthol concentration:</td>
<td></td>
</tr>
<tr>
<td>• Maximum VAS cold, unpleasantness, pain</td>
<td><em>Pearson’s or Spearman’s Correlation Coefficient</em></td>
</tr>
<tr>
<td>• PRI and NWC</td>
<td></td>
</tr>
<tr>
<td>4. Association between CPT and response at each menthol concentration:</td>
<td></td>
</tr>
<tr>
<td>• <em>Forearm CPT (°C)</em></td>
<td><em>Pearson’s or Spearman’s Correlation Coefficient</em></td>
</tr>
<tr>
<td>• Maximum VAS cold, unpleasantness and pain</td>
<td></td>
</tr>
<tr>
<td>5. Differences in intensity and quality of response at each concentration between those with CPT&lt;15°C and those with CPT&gt;15°C:</td>
<td></td>
</tr>
<tr>
<td>• Maximum VAS cold, unpleasantness, pain x3</td>
<td><em>Independent t-tests or</em></td>
</tr>
<tr>
<td>• PRI and NWC x3</td>
<td><em>Mann-Whitney U tests</em></td>
</tr>
<tr>
<td>• Time to onset x3</td>
<td></td>
</tr>
</tbody>
</table>

*Analyses were completed on Liquids A, B & C only since initial analysis of control Liquid D showed no response for cold, unpleasantness or pain ratings. Inclusion of Liquid D in the repeated measures ANOVA may therefore have distorted the final ANOVA statistic.*
3.5 Results

- **Cohort characteristics**

Thirty-two subjects participated in this study, seven males and 25 females. The mean age of the group was 29 years (range 19-61). Differences in measures between gender and age response were not calculated in this study.

<table>
<thead>
<tr>
<th>Age (range) (years)</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;20</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>20-29</td>
<td>16</td>
<td>19</td>
<td>35</td>
</tr>
<tr>
<td>30-39</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>40-49</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>50-59</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>60-69</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><strong>7</strong></td>
<td><strong>25</strong></td>
<td><strong>32</strong></td>
</tr>
</tbody>
</table>

**Hypothesis 1**

*There will be significant dose-dependent differences in intensity and quality of sensation experienced during the 15-minute application of liquid formulations containing different concentrations of menthol.*

- **Intensity:** there was a significant concentration-dependent difference in VAS intensity as measured by area under the VAS-time curve and by maximum VAS for cold, unpleasantness and pain.

Ethanol control (Liquid D) did not evoke any sensations of cold, unpleasantness or pain in any subject. Further between-concentration analyses were therefore limited to liquid formulations A, B and C.

There was a significant effect of concentration for total VAS intensity (AUC): VAS cold $F_{(2,62)}= 11.31$, $p<.001$; VAS unpleasantness $\chi^2(2)= 11.58$, $p=.003$; VAS pain $\chi^2(2)= 16.44$, $p<.001$. The increase in menthol concentration from Liquid A to C was mirrored by the gradual increase in intensity (Table 3.3). Post hoc contrasts showed that the difference was significant between each concentration for cold VAS (Liquids A-B $p=.017$; B-C $p=.019$; A-C $p<.001$). For both unpleasantness and pain VAS there were significant differences between the lower concentrations of A and B (unpleasantness $p=.012$; pain
p=.002) and between Liquids A and C (unpleasantness p<.001; pain p=.003) but not between Liquids B and C (unpleasantness p= .108; pain p=.717).

**Table 3.3:** Mean (SEM) area under the time-VAS curve values for each sensation at each concentration

<table>
<thead>
<tr>
<th>Sensation</th>
<th>Liquid A Mean</th>
<th>Liquid A SEM</th>
<th>Liquid B Mean</th>
<th>Liquid B SEM</th>
<th>Liquid C Mean</th>
<th>Liquid C SEM</th>
<th>F(2,62)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold</td>
<td>273.8</td>
<td>36.2</td>
<td>337.9</td>
<td>32.7</td>
<td>428.5</td>
<td>45.5</td>
<td>11.31</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>Unpl</td>
<td>79.7</td>
<td>17.7</td>
<td>151.5</td>
<td>31.2</td>
<td>199.3</td>
<td>30.3</td>
<td>11.58</td>
<td>.003*</td>
</tr>
<tr>
<td>Pain</td>
<td>26.7</td>
<td>10.5</td>
<td>91.7</td>
<td>24.0</td>
<td>91.6</td>
<td>29.4</td>
<td>16.44</td>
<td>&lt;.001**</td>
</tr>
</tbody>
</table>

Maximum VAS intensity experienced during the 15-minute application also showed a significant concentration-dependent effect for each sensation: VAS cold $F_{(2,62)}= 8.67$, p<.001; VAS unpleasantness $\chi^2(z)= 14.14$, p=.001; VAS pain $\chi^2(z)= 11.74$, p=.003 (Figure 3.1). Post-hoc contrasts showed a similar pattern to AUC analysis, with significant differences for all sensation between concentrations A and C. Significant differences in unpleasantness and pain were found between concentrations A and B but not between the higher concentrations of B and C (unpleasantness p=.086; pain p=.710). Maximum cold VAS however showed a significant difference between Liquids B and C but not between A and B (p=.322).

![Graph showing VAS intensity ratings](image)

**Figure 3.1:** Differences in maximum VAS intensity ratings between menthol concentrations (Liquids A, B and C).

- **Intensity:** there was a dose-dependent effect in the percentage of subjects reporting unpleasantness and pain.
All menthol concentrations evoked cold, unpleasantness and pain VAS >0 in a proportion of healthy subjects: cold VAS >0 by 94-100% subjects; unpleasantness VAS >0 by 66-84% subjects; pain VAS >0 by 35-47%. This was dose-dependent for unpleasantness and pain VAS. However, mean VAS intensity was low, ranging from 30.0-39.9/100 for cold, 11.3-25.1/100 for unpleasantness and just over 4-12.2/100 for pain.

**Intensity:** there was good intra-subject consistency between menthol concentrations

There were strong significant positive correlations for maximum cold VAS between concentrations: range r=.773, p<.001 to r=.708, p<.001. Correlations between concentrations for unpleasantness intensity were moderate: range r=.551, p=.001 to r=.495, p=.004. Those for VAS pain were most variable although still significant: (r=.507, p=.003 to r=.360, p=.043).

**Quality:** There was a significant concentration-dependent difference in word choice, as measured with the PRI and NWC indices.

There was a significant concentration-dependent effect for descriptor choice: PRI $F_{(2,62)}=26.33$, p<.001; NWC $F_{(2,62)}=19.62$, p<.001. Higher scores for PRI and NWC were recorded at higher menthol concentrations. Post-hoc contrasts showed that the difference was significant between the lowest concentrations of Liquids A and B (PRI: $t_{(31)}=-7.01$, p<.001; NWC: $t_{(31)}=-5.80$, p<.001), and also between Liquids A and C (PRI: $t_{(31)}=-5.56$, p<.001; NWC: $t_{(31)}=-5.40$, p<.001). However, in similar manner to VAS intensity ratings, there was no significant difference in quality scores between the two highest concentrations of Liquids B and C (PRI: $t_{(31)}=-1.14$, p=.265; NWC: $t_{(31)}=-1.13$, p=.266).
• **Quality:** Higher concentrations of menthol provoked selection of more intense words

When word choice was analysed, a dose-dependent effect was also found (Figure 3.4). During liquid A application, subjects mostly selected relatively mild words: cool (90.6%); tingling (43.8%). At the higher concentrations, cold or icy were more often chosen (Liquid C: cold 93.8%; icy 40.6%) plus unpleasant words such as burning (Liquid B 43.8%; Liquid C 59.4%), stinging (53% for both B and C) and prickling (25% and 21.9%).

**Figure 3.4:** Percentage of subjects choosing the most commonly selected words for each liquid.

• **Quality:** there was a difference between concentrations in the percentage of time types of words were chosen.

Cold-type words (see classification in Figure 3.5) were used for the highest proportion of the time (56%) during the lowest concentration Liquid A and progressively less selected as the concentration rose (40%, 37%). In contrast, heat-type words (including burning) rose in percentage use as the concentration increased (13%, 22%, 24%). Dysaesthetic words, both mild and intense, featured for slightly more of the time for the highest concentration.
Figure 3.5: Percentage use of words types during 15-minutes application of each concentration.

**Word classification:**
- **Cold words**
  - cool, cold, icy/freezing
- **Heat words**
  - warm, hot, burning
- **Intense dysesthetic**
  - stinging, pricking
- **Mild dysesthetic**
  - tingling, numb, itchy
- **Penetrating words**
  - spreading, penetrating
- **Other words**
  - aching, sharp, intense

- **Quality:** there was good intra-subject consistency between concentrations for PRI and NWC descriptor indices

PRI showed good correlation between Liquids A and B (Figure 3.6a) and between Liquids B and C (Figure 3.6b), although there was no significant correlation between Liquids A and C (r=.216, p=.232). NWC index showed the same pattern: good correlations between Liquids A and B (r=.713, p<.001) and between B and C (r=.499, p=.004) but less so between Liquids A and C (r=.403, p=.022).

Figure 3.6a-b: Correlations between PRI score at: a) Liquids B and C; b) Liquids A and B
Hypothesis 2

There will be significant dose-dependent differences in temporal characteristics between menthol formulations: time to onset of first VAS rating, time to maximum rating and overall time-course of sensation intensity and quality.

- There was a significant dose-dependent difference in time to first VAS rating

Time to onset of first cold rating >0 showed a significant dose-dependent effect (F(2,62)= 13.28, p<.001), with higher concentrations of menthol provoking earlier onset.

Time to onset for pain VAS also showed a significant concentration-dependent effect ($\chi^2(3)$= 7.91, p=.019) however, unpleasantness did not ($\chi^2(3)$= 5.29, p=.071) (Figure 3.7). Post hoc contrasts showed that the difference in time to onset was significant between each concentration for cold (concentrations A-B t(31)= -2.27, p=.032; B-C t(31)= -3.32, p=.003; A-C t(31)= -4.70, p=<.001). For both unpleasantness and pain, post-hoc Wilcoxon signed ranks tests showed a significant difference only between concentrations A and C (unpleasantness p=.034; pain p=.047).

![Figure 3.7: Time to onset of first recorded sensation for each concentration.](image)

There was a clear dose-dependent pattern to timing of onset for each sensation: cold was experienced significantly sooner than unpleasantness or pain for all three concentrations (Figure 3.8).

![Figure 3.8: Time to onset of first recorded sensation for each concentration.](image)

- There was no difference in time to maximum VAS rating for any sensation

93
There was no significant concentration-dependent effect in terms of differences for time to peak VAS for cold ($F_{(2,62)} = .991$, $p = .360$) or for unpleasantness ($\chi^2(2) = 1.87$, $p = .393$) or pain ($\chi^2(2) = .083$, $p = .959$). Cold peaked at between 9 and 10 minutes, pain at between 10 and 10.5 minutes and unpleasantness at between 10.5 to 11 minutes.

![Figure 3.9: Time to peak (maximum) VAS of each sensation at each concentration](image)

- A similar but dose-dependent time-course pattern was shown between concentrations for cold, pleasantness and pain

Time course followed a similar pattern both between concentrations for each VAS sensation intensity and also between sensations (Figure 3.10a-c). For cold and unpleasantness at each concentration, there was a steady rise in intensity to a peak just before removal of the stimulus. Removal was immediately followed by a sharp decline in sensation. Pain VAS showed a less markedly sharp rise due to low VAS ratings.

![Figure 3.10a-c: Mean VAS intensity at each minute during testing for each sensation at each concentration: a) cold; b) unpleasantness; c) pain](image)
• A dose-dependent time-course pattern was also shown between high and low concentrations for most frequently selected words

When time course of the most frequently selected words was considered, there were clear differences between Liquid A formulation and the two higher concentrations. Figures 3.11a-b illustrate the timing differences between Liquid A and Liquid C for key words. During application of higher concentration Liquid C, cold was reported by more than 50% of subjects within the first 2 minutes, peaking at 7 minutes and then gradually declining. For lower concentration Liquid A, cold was experienced by fewer subjects and only gradually increased to its peak, also around 7 minutes. There was also a difference in temporal experience of burning and stinging. For Liquid C, both sensation qualities started to be reported within the first 2 minutes, steadily rising to a peak at 10-11 minutes. During Liquid A those fewer subjects who reported any burning or stinging did so within the first 4 minutes.

![Figure 3.11a: Time-course for percentage of subjects selecting key words during Liquid C](image1)

![Figure 3.11b: Time-course for percentage of subjects selecting key words during Liquid A](image2)
Hypothesis 3

There will be a significant association between measures of sensation intensity and measures of sensation quality at each concentration.

- There were some moderate associations between VAS intensity and PRI index

Cold VAS correlated moderately with PRI for Liquids A and B (Liquid A: r = .419, p = .017; Liquid B: r = .469, p = .007). There was no correlation between VAS cold and PRI for the highest menthol concentration (Liquid C), yet unpleasantness was positively correlated with PRI only for Liquid C (r = .471, p = .006). There were no correlations between VAS pain and PRI at any concentration.

To explore this further, subjects were grouped dichotomously according to high or low PRI scores (> or < 1SD from the mean) to determine if there were group differences in VAS values. T-tests showed there was a significant difference in cold and unpleasantness VAS for Liquid C (Table 3.4).

Table 3.4: During Liquid C, subjects in the high PRI score group (score > mean +1SD) reported significantly higher VAS intensity for cold and unpleasantness, but not pain.

<table>
<thead>
<tr>
<th></th>
<th>Low PRI (n=23)</th>
<th>SEM</th>
<th>High PRI (n=9)</th>
<th>SEM</th>
<th>t(31)</th>
<th>p=</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS (0-100) Cold</td>
<td>34.8</td>
<td>3.7</td>
<td>53</td>
<td>8.2</td>
<td>-2.34</td>
<td>.026*</td>
</tr>
<tr>
<td>Unpl</td>
<td>18</td>
<td>3.2</td>
<td>43.3</td>
<td>6.3</td>
<td>-4.00</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>Pain</td>
<td>8.3</td>
<td>2.6</td>
<td>18.9</td>
<td>9.0</td>
<td>-1.52</td>
<td>.138</td>
</tr>
</tbody>
</table>

Hypothesis 4

There will be a significant association between cold pain threshold temperature and sensory response to each menthol formulation.

- CPT and menthol intensity: There was a significant correlation between CPT temperature and VAS pain rating for the highest menthol concentration only

Mean forearm CPT was 8.95°C (standard deviation 6.4, range 0-21.3°C). Only very limited significant correlations between CPT and VAS intensity ratings were seen, with a moderate positive correlation shown only between CPT and the highest concentration (Liquid C) for VAS pain intensity: r = .484, p = .008. There were no other correlations between CPT and any other VAS sensation at any concentration.

- CPT and menthol quality: There were significant correlations between CPT temperature and PRI and NWC descriptor index values
There were significant moderate positive correlations between CPT temperature and scores for PRI and NWC for Liquids A and B, but no correlation between CPT and the highest concentration Liquid C (Table 3.5).

**Table 3.5:** Correlations between CPT (°C) and PRI and NWC descriptor indices at each concentration

<table>
<thead>
<tr>
<th></th>
<th>Liquid A</th>
<th></th>
<th>Liquid B</th>
<th></th>
<th>Liquid C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRI</td>
<td>NWC</td>
<td>PRI</td>
<td>NWC</td>
<td>PRI</td>
<td>NWC</td>
</tr>
<tr>
<td>CPT</td>
<td>r=.494</td>
<td>r=.423</td>
<td>r=.493</td>
<td>r=.358</td>
<td>r=.028</td>
<td>r=.057</td>
</tr>
<tr>
<td></td>
<td>p=.004*</td>
<td>p=.016*</td>
<td>p=.004*</td>
<td>p=.044*</td>
<td>p=.879</td>
<td>p=.755</td>
</tr>
</tbody>
</table>

**Hypothesis 5**

*Individuals with an elevated cold pain threshold (CPT>15°C) will exhibit an elevated response and different time to onset for each menthol formulation when compared with those with CPT<15°C.*

- **Intensity:** There were trends towards those with high CPT having higher VAS ratings at all concentrations, but this only reached significance for the higher concentrations.

When subjects were divided according to CPT cut-off of 15°C, 9 subjects (28%) exhibited CPT >15°C. There were no significant differences between CPT groups for total VAS intensity (AUC) for any sensation or any liquid formulation (p=.106 to p=.928). For maximum VAS during application (Figure 3.12), there were no significant differences during application of Liquid A. For Liquids B and C there were trends towards those with high CPT reporting higher VAS intensity, but this only reached significance during Liquid B for cold (p=.025) and unpleasantness (p=.050). For Liquid C, only pain VAS was significantly higher for the high CPT group (p=.022) although VAS cold was close to significance (p=.068)
Figure 3.12: Differences in maximum VAS intensity ratings for subjects in high and low CPT groups

- **Quality**: There were trends towards those with high CPT having higher PRI scores, although this was only significant for Liquid B. There was minimal difference in NWC values between concentrations.

There was a significant difference between CPT groups in PRI score during Liquid B ($t_{31} = -2.76$, $p = .013$) but not significant difference for Liquid A ($t_{31} = -1.53$, $p = .136$) or Liquid C ($t_{31} = -1.12$, $p = .263$). There were no significant differences between groups for NWC index during any of the liquid formulations (Figure 3.13).

Figure 3.13: Differences in PRI scores for subjects in high and low CPT groups

- **Quality**: Those with high CPT consistently selected burning more often than those with low CPT, at all concentrations. Liquid B was most discriminatory concentration of the three (Figure 3.14a-c).

For all concentrations, the descriptors cool, cold and icy/freezing did not discriminate between high and low CPT. For the lowest concentration, Liquid A, burning, stinging, and tingling most clearly identified CPT group membership. The same was true for Liquid B, with penetrating and prickling also showing clear difference between groups. For Liquid C, burning was the only word that differentiated those with high CPT.
**Time to onset**: Those with high CPT reported earlier onset for each sensation

There were observable differences between groups in time to onset (Figure 3.15). Those with high CPT experienced each of the VAS sensations at each concentration between 1 minute (for cold VAS) and 3 minutes (for pain VAS) earlier than those with low CPT. However, the difference was only significant for Liquid B at each of the VAS sensations.

![Figure 3.14a-c: Comparison of most frequently chosen descriptors between those with high or low CPT values for all three concentrations.](image)

![Figure 3.15: Comparison between CPT < or >15°C groups for time to onset of VAS cold, unpleasantness and pain at each concentration](image)
• **ROC curves showed good sensitivity and specificity for Liquid B as a predictor of CPT group.**

Receiver operating characteristic (ROC) curves were also applied to assess the sensitivity and specificity of VAS and PRI variables in predicting membership of CPT<15°C or CPT>15°C groups. The best results were seen for Liquid B for VAS cold and PRI. For cold VAS, a cut-off at 21/100 resulted in an area under the curve of .717, sensitivity of 89% but lower specificity of 65%. PRI showed similar results: with a cut-off of 10.5, an area under the curve of .710, sensitivity was 89% and specificity 61%.

### Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Accepted/Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. There will be significant dose-dependent differences in intensity and quality of sensation.</td>
<td>Accepted</td>
</tr>
<tr>
<td>2. There will be significant dose-dependent differences in temporal characteristics between menthol concentrations.</td>
<td>Accepted</td>
</tr>
<tr>
<td>3. There will be a significant association between sensation intensity and sensation quality at each concentration.</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>4. There will be a significant association between CPT temperature and sensory response to each menthol concentration.</td>
<td>Accepted</td>
</tr>
<tr>
<td>5. Those with CPT&gt;15°C will exhibit a significantly elevated response and different time to onset for each menthol concentration.</td>
<td>Partially accepted</td>
</tr>
</tbody>
</table>
3.6 Discussion
This study sought to investigate systematically whether different concentrations of topical menthol evoked dose-dependent differences in intensity of sensation or quality of sensation. The potential validity of a menthol test for assessment of cold hyperalgesia was also investigated.

**Dose dependent effects**
Preliminary laboratory studies had shown that graded concentrations of menthol in a simple liquid vehicle showed dose-dependent release across a semi-permeable membrane. The current study found the same dose-dependent pattern when the sensory effect of three concentrations of menthol in solution applied to the forearm of healthy subjects was evaluated. Subjects reported higher levels of intensity and changes in quality for the higher concentrations of menthol. Times to onset of first cold, unpleasantness and pain sensations were also earlier for the higher concentrations. This clearly shows that different concentrations of menthol are graded in the sensory system in a similar manner to different cold temperatures. Liquid A elicited a predominantly cool or cold sensation, whereas Liquids B and C evoked more intense cold or icy sensations, associated with noxious sensations such as stinging or burning.

In terms of mechanisms, these sensory effects indicate that, in normal subjects, lower concentrations of menthol may predominantly activate TRPM8 channels expressed on thermal Aδ-fibres, which signal non-noxious, lower intensity cold. Some activation of TRPM8 receptors on nociceptive Aδ-fibres is also indicated, even during low concentrations, since some individuals reported cold associated with dysaesthetic sensations such as tingling or stinging, which was described as unpleasant, although at a very low level of intensity. At higher concentrations, TRPM8 receptors appear to still have been activated, as evidenced by the continued report of cold, albeit at higher intensity. In a non-pathological environment, increased intensity of cold may reflect the gradual binding of all four menthol sites on each TRPM8 channel (Janssens & Voets, 2011) or it may reflect the greater quantity of menthol molecules available to bind to TRPM8 channels. Higher concentrations of menthol were certainly able to permeate the skin as efficiently as lower concentrations. Indeed the finding that time to onset of first cold sensation was quicker for higher concentrations (Figure 3.7) suggests that increased volume diffused more efficiently. This is supported by pharmacological studies which use menthol as a penetration enhancer. These studies have reported that
penetration of a drug is increased in a linear manner as menthol concentration increases (Fang et al., 2008).

In addition to more intense cold sensations, at higher concentrations participants were more likely to report noxious sensations. The cold temperature quality was more likely to be described as icy/freezing and there was a more frequent report of stinging, prickling and tingling sensations, associated with higher VAS intensity ratings for unpleasantness and pain. Dysesthetic or noxious sensation is associated with activation of nociceptors. TRPM8 is reported to be expressed on both Aδ-thermo fibres and also on cold and warm sensitive Aδ-nociceptive fibres, identified by low level co-expression of Nav1.8 and TRPA1 / TRV1 channels (Akopian, 2011; McCoy et al., 2011). It has been proposed that these Aδ nociceptors continue to signal cold but may also signal low intensity noxious sensations (Campero et al., 2009), acting as part of an 'early warning' system. As levels of available menthol increases still further, cold and heat sensitive higher threshold ‘c2-fibres' may well become activated. These fibres express some TRPM8 channels but predominantly express nociceptive marker channels such as TRPA1, TRPV1 and Nav1.8 (Campero & Bostock, 2010). Consequently their activation is associated with more unpleasant sensations. Report of paradoxical burning sensation is more difficult to explain with certainty, as there are a number of different theories regarding mechanism in both normal and pathological examples. Burning may simply imply activation of nociceptors that express TRPV1 channels, such as for c2-fibres. However, c2-fibres are associated with lower intensity response (Campero & Bostock, 2010) whereas in the current study choice of burning tended to be associated with higher levels of unpleasantness and pain. This may suggest that for some individuals, higher concentrations of menthol activated additional polymodal high threshold c-fibres, which signal more intense nociceptive qualities via activation of TRPA1 and TRPV1 channels (Okazawa et al., 2004; Kobayashi et al., 2005). The fact that unpleasantness and then pain were reported consistently later than cold supports the hypothesis that slower c-fibres were responsible.

**Characterisation of normal response to menthol**

The majority of subjects reported that menthol at all concentrations evoked a predominantly cold sensation, which increased in intensity with increasing concentration. Some degree of low intensity unpleasantness was also experienced by more than 50% of subjects even during lower concentration Liquids A and B. In contrast to the majority, two subjects reported no cold sensation at all during higher
concentration menthol applications (Figure 3.2). One subject (subject 8) consistently reported burning and stinging sensations, associated with soreness for both Liquid B and C. The other (subject 19) reported warmth, with tingling and some prickling but no soreness just during Liquid B, although during Liquid C these same sensations were also associated with cold. Both were clearly atypical (abnormal) responses may provide an interesting insight into how cold is signalled, although recent progress in understanding about the variety of fibres that co-express moderate or small numbers of TRPM8 channels with Nav1.8 or TRPV1 or TRPA1 means that it is difficult to be conclusive. Lack of any cold sensation strongly suggests that TRPM8-expressing Aδ thermal fibres were not activated (Knowlton et al., 2013). However, low threshold cold and warm sensitive (TRPM8- and TRPV1 co-expressing) Aδ nociceptors or c2-fibres (Campero & Bostock, 2010) may instead have been activated, particularly for subject 19. In this case menthol would bind to TRPM8 and initiate intra-cellular cascades to activate calcium influx, but the fibre signalled a nociceptive response associated with warmth and mildly unpleasant tingling and prickling. This subject did not report any sensation until four minutes into the application, which may suggest c-fibre speed. Subject 8 exhibited a much more intense response suggesting that menthol triggered a full-blown nociceptive response from normally high threshold c-fibres. It may be that TRPA1 channels were directly activated by menthol, as proposed by Xiao et al. (2008), although it is suggested that this only occurs when TRPA1 thresholds are already lowered by some other mechanism (Bautista et al., 2012). It is unlikely that Aδ-fibre TRPM8 channels were not activated by circulating menthol, so cold signals from may well have been inhibited at central terminals. This may have ben locally-driven by excessive c-fibre barrage: Fruhstorfer (1984) demonstrated that A-fibre block causes a cool sensation to be experienced as noxious and icy, burning and stinging and Susser et al. (1998) demonstrated that this sensation is conducted via the slowest c-fibres. Alternatively altered centrally-driven disinhibition may be responsible, since paradoxical burning in response to a cold stimulus has been reported as indicative of altered spinal inhibitory drive (Leung et al., 2001; Lindstedt et al., 2011).

There is almost no data with which to compare these results. Green and Schoen (2007) applied 10% l-menthol in a similar aqueous-ethanol vehicle to a slightly larger area on the volar forearm before assessing cold and nociceptive sensations. Baseline menthol effect was rated as 'between barely detectable and weak', with cold being felt marginally more than pain. Hatem et al. (2006) used a higher concentration of menthol as an experimental cold sensitising agent. The volume and area of application was very
similar to the current study. However when subjects were asked for spontaneous words to describe what they were feeling the study reports that 90% of subjects reported spontaneous coolness after 10 minutes, 10% reported warmth and no-one reported any pain. No further details are provided. The higher reports of unpleasantness and pain from subjects in the current study are most likely to reflect difference in method. Hatem et al. (2006) did not use a systematic questioning approach since the study focus was on the effects on cold temperature and pain perception and there was little interest in the direct sensory response to menthol. In contrast the current study required participants to concentrate on the sensation experienced within the context of specific VAS scales and descriptor lists. It may therefore be that the current study was biased towards greater report of sensory effect.

However, additional analysis of construct and criterion validity suggested that the current study was valid. There was good evidence for construct validity with response consistent between concentrations. Strong to moderate significant positive correlations were shown between concentrations for each VAS sensation (range r=.773 to r=.495). There was also an association between the domains of intensity and quality, mainly between PRI and unpleasantness intensity (Table 3.4).

There was also evidence for criterion validity in this data. Using CPT>15°C as the cut-off for classification into cold hyperalgesic and non-hyperalgesic groups, those with high CPT showed signs of significantly elevated intensity and quality ratings, predominantly for Liquid B application. Cold and unpleasantness were significantly higher for the cold hyperalgesic group during Liquid B (p=.025 and p=.050 respectively) and PRI scores for Liquid B were also significant higher (p=.013). Figures 3.14a-c illustrate that at all concentrations those with higher CPT more frequently reported burning and stinging sensations whereas those with lower CPT reported more tingling. It is also notable that those with CPT-assessed cold hyperalgesia also reported significantly earlier time to onset for all sensations during Liquid B (Figure 3.13). These results indicate a clear association between conventionally measured cold hyperalgesia and application of the middle concentration of menthol solution (Liquid B) in this study, suggesting that this concentration may be the most discriminatory for identification of cold hyperalgesia with menthol. Chapter 2: 2.5 of this investigation reviewed the likely mechanisms that may drive cold hyperalgesia. The earlier initiation of response, higher levels of unpleasantness and more frequent sensations of burning and stinging shown in participants with CPT-cold hyperalgesia implies that similar mechanisms may drive an
atypical response to menthol and emphasises the close relationship between the two cold modalities. Final ROC curve analysis also showed that Liquid B was best able to predict CPT group membership. Although not unequivocal, these results are a further indication that a consistent menthol stimulus may evoke responses that are an equally valid measure of cold hyperalgesia.

This study provided evidence to support in-principle content and construct validity for the menthol test. The discussion above underlines the benefits of a measurement system that provides information from more than one sensory domain. When trying to identify cold hyperalgesia, having data for both intensity and quality of a sensation enables more comprehensive characterisation to occur than is possible with a single outcome variable, such as with CPT temperature. Once the specifics of which sensation type should be included in the VAS element and which words should be included in the final descriptor list, there is potential for improved content validity in a test with a scoring system that combines both types of data.

Assessment of test procedure, practical and safety issues:

In addition to statistical analysis of the sensations provoked by different concentrations, the liquid menthol test was also evaluated for its safety and practicality.

- **Adverse responses**
  Approximately 40% subjects experienced visible erythema on removal of the liquid formulation at 15 minutes. In half of these participants the erythema was quite vividly red and had spread beyond the limits of the gauze, most often onto the ulnar aspect of the forearm. This appeared to be associated with leakage and pooling of the liquid as discussed below. Once the menthol liquid was washed off, the majority of subjects reported no discomfort from the skin reddening and by 30 minutes (15 minutes post menthol removal) the erythema had dispersed.

- **Test practicalities**
  The lack of viscosity of the liquid solutions was a problem for some individuals. Subjects were required to maintain their forearm in near full supination in order to expose the volar surface. Many participants found this a challenging position to maintain for 15 minutes and so, as their forearm relaxed into a more neutral position, the liquid formulation tended to pool on the ulnar side. Pooling was directly related to an increase in more intense sensations such as stinging and spreading since the more intense area seemed to be on the ulnar side. Experimental reduction in the volume of liquid with a small number of additional subjects made no difference to this effect. It
was therefore decided that, although the basic menthol solution was sound as far as
dose-dependent effect was concerned, greater consistency and reliability would be
achieved with a more viscous delivery system.

The ultimate goal is for a topical menthol test to be used in a clinical setting. From a
positive perspective, first indications are that the time-course for menthol release is
suitable for a clinical environment, with peak sensation occurring at between 9 and
10.5 minutes. Using VAS scales and descriptor indices to quantify response appears to
provide appropriate data. However further refinement will be needed, both in terms of
content and delivery, to generate a tool that is practical to apply, easy to interpret and
which provides appropriately sensitive and specific data for identification of cold
hyperalgesic individuals.
3.7 Summary

This first study found clear dose-dependent effects for intensity, type of sensation and release time between liquid formulations of menthol at low medium and high concentrations. Menthol evoked cold sensations in 95% of healthy subjects, with additional low-intensity dysesthetic sensations experienced by approximately ⅕ of subjects. These findings suggest that increasing concentrations of menthol activate similar mechanism as described for decreasing cold temperature and support the basic science evidence that as concentrations increase, the ratio of activated cold-sensitive Aδ to c-fibres is reversed, resulting in greater unpleasantness and changed quality of sensation.

A small percentage of subjects reported no cold but only noxious burning and stinging sensations. This suggests that a menthol concentration that normally evokes non-noxious sensations in most healthy subjects is able to clearly differentiate those with atypical responses by the lack of TRPM8 effect. Liquid B (middle concentration) appeared to be the most effective at differentiating between high and low responses.

Even though there were some issues with pooling of the liquid, possibly resulting in higher than expected reports of unpleasant sensations such as stinging, the study showed good signs of internal content and construct validity. There were clear correlations between concentrations for each VAS sensation and PRI score and some association between descriptor score and VAS unpleasantness rating. There were also clear indications of criterion validity, with Liquid B in particular differentiating between those with and without thermode-cold hyperalgesia.

This first study therefore proved good indications that a more viscous formulation would have the capacity to identify cold hyperalgesic individuals.
Chapter 4

Study Two

Evaluation of the gel test formulation

4.1 Abstract

Background and Aims

Chapter 3 found that, although a liquid formulation of menthol showed clear dose-dependent effects for intensity and quality of response, the solution provided a sensory stimulus that was less consistent than desired. This study used the same methodology to evaluate the dose-dependent response of two menthol gel formulations in healthy subjects. In addition the study aimed to evaluate whether either of these gels was able to clearly identify an abnormal response and which concentration was most effective.

Method

The same blinded, randomised, repeated-measures design was used as for Study 1, with each subject experiencing both of the menthol gels on separate occasions at the same volar forearm site for 15 minutes. A very similar response measurement method was used, with heat VAS added to the intensity ratings. VAS ratings were taken every one-minute and descriptors chosen from an MPQ descriptors list every two minutes.

Results

A significant dose-dependent effect was seen for both intensity and quality of sensation. There were significant dose-dependent effects for VAS cold (p = .013), heat (p = .001), unpleasantness (p = .002) and pain (p = .001). There were also significant dose-dependent effects for PRI and NWC descriptor indices (p = .002 and p = .006). There were strong similarities between responses to gel and liquid formulations. An abnormal response was best demonstrated by presence of high unpleasantness and pain VAS combined with either high heat or cold intensity and high descriptor score. Inclusion of the specific words icy, burning and stinging was also required in an abnormal response for healthy subjects. Gel B was better able to discriminate these characteristics than the lower concentration Gel A.

Conclusion

Menthol gel formulations evoked similar dose-dependent responses to equivalent liquid formulations and showed similar internal construct validity. Gel formulation B was able to discriminate most clearly between those with normal and abnormal responses.
4.2 Introduction & Background

The first study in this series (Chapter 3) found that a simple topical menthol solution showed good concentration-dependent effects when applied to healthy subjects. Higher concentrations of menthol evoked higher intensities of cold, unpleasantness and pain, which was felt increasingly as an icy/freezing, burning or stinging sensation. However, the low viscosity of this liquid formulation resulted in inconsistent application to the skin. As a result the normal response was difficult to characterise reliably. The current study therefore sought to assess whether a menthol formulation to which a gelling agent had been added showed similar dose-dependent characteristics.

Laboratory studies that were carried out to determine the optimal amount of a gelling agent that was needed to create a formulation sufficiently viscous to avoid pooling whilst not preventing the active menthol from permeating the skin at the same rate as the solution. A gelling agent was selected that is widely available and inexpensive, important factors for a low-cost commercial test. The particular gelling agent also has good solubility and has shown good bio-adhesion, whilst maintaining low skin irritation effects. However, any gelling agent, if used in too high a concentration, may slow the diffusion of menthol through the formulation and so slow its release and subsequent permeation through the skin (Stamatialis et al., 2004). It was therefore important to assess in vivo whether there were any changes in the response characteristics of the menthol formulation with the addition of menthol.

Since the primary goal for the current study was to compare systematically equivalent gel and previous liquid formulations, and since the measurement method had shown good content and construct validity in Chapter 3, few changes were made. An additional VAS scale for heat intensity was added to the existing three VAS (cold, unpleasantness and pain) since words relating to warmth or heat had been selected more often than anticipated in the first study. Paradoxically, concurrent sensations of warmth and cold during a cold stimulus has been described both as a normal response (Greenspan et al., 1993) and as a sign of a ‘malfuctioning’ thermal system (Susser et al., 1998; Green et al., 2008). For example, Susser (1998) used QST and nerve conduction studies to find that sensations of paradoxical heat were conducted by nerves with a velocity very similar to that for c-fibre heat transmission and considerably slower than that for a cold sensation. The authors concluded that a malfunctioning of the cold-thermal pathway was the cause. Green (2008) also attributed the abnormal heat sensations evoked by non-noxious cooling of certain
areas in the skin to an abnormal inhibitory system. Clearly defined measurement of paradoxical heat during a cold sensation is therefore an important factor in the assessment of cold hyperalgesia as it may point to dysfunctional inhibitory mechanisms and would potentially add useful psychophysical data to the current discussions about cold and cold pain transduction.

Also in response to findings from Study 1 (Chapter 3), it was decided to limit testing of menthol gel to two concentrations since the ultimate goal was to determine a single final concentration. Study 1 indicated that the highest concentration (Liquid C) was the least useful of the three formulations. For many of the intensity and quality variables, Liquid C produced no additional effects beyond that for Liquid B. No significant differences were found between Liquid B and C for unpleasantness or pain VAS, or for PRI or NWC indices. In contrast significant differences in response between Liquids A and B were consistently found for both VAS and descriptor variables. It was therefore decided to evaluate a gel formulation of concentrations A and B in the current study.

Determination of the most appropriate stimulus intensity is a vital aspect in the development of a valid menthol test designed to identify individuals with cold hyperalgesia. Recent new developments of clinical cold hyperalgesia/allodynia tests have taken several approaches. Maxwell and Sterling (2012) have recently reported evaluation of an ice-based contact cold test to be used in combination with an NRS pain scale to assess cold response in individuals with whiplash associated disorder. However, the problem with this test, simple and inexpensive though it is, relates to the type of stimulus. Ice by definition must be a supra-threshold cold stimulus for many pain-free individuals since the temperature of an ice block is approximately 2°C to 4°C. As the authors themselves acknowledged, subjects with cold hyperalgesia were unable to manage the contact cold for even the 10 seconds of the test, resulting in issues with ceiling effects from many pain-free and healthy subjects. Ideally a test for hyperalgesia involves a stimulus that is not noxious to many individuals so that a range of data is produced. A test for cold alldynia certainly must involve a sub-threshold stimulus so, in contrast, several groups have developed simple ‘bedside’ tests for neuropathic-type pain which includes an assessment of response to cold. Most commonly this has involved using an object assumed to be at a temperature of around 20°C: for example a brass cylinder (Nijs et al., 2010; Eaton et al., 2012) or a glass with water in it (Baron et al., 2010). This test could be sufficient to differentiate those with severe cold alldynia indicative of nerve injury, when used alongside additional tests of neurological
function. However, within the context of identifying those with less extreme cold hyperalgesia that may be amenable to appropriate intervention, the test has too much potential for variability, and lacks precision and reliability. The menthol test has potential advantages over both of these recent methods because it is not dependent on thermal mechanisms to evoke a cold response. Menthol is capable of directly activating TRPM8, a key cold sensing receptor located in the skin (McKemy, 2005). Chapter 3 indicated that varying concentrations of menthol evoke varying intensities of cold and noxious responses. The goal is to find a concentration that is non-noxious / sub-threshold for the majority of healthy individuals but that evokes a clearly noxious response in those with pain processing that has developed an element of central augmentation.

The recent studies attempting to find simple clinical solutions to the problem of cold pain testing illustrate that it is increasingly seen as an important feature of chronic pain that needs to be assessed. Additional features of these suggested clinical cold pain tests include the fact that they do not involve expensive equipment and take very little time to apply. The problem with all of these suggested tests is that they are neither reliable nor valid and so either will create too many false positive or too many false negative results. Consequently, the current will focus on reviewing content and construct validity for the gel formulations of the menthol test. Once validity has been ascertained, future studies will concentrate on improving some of the current practical difficulties that would limit the application of the test in the clinical environment.

The primary aim of this study therefore was to determine whether a gel formulation using menthol concentrations A and B from Study 1 showed similar dose-dependent effects for intensity and quality of sensation and response time as had the liquid formulation. The study also sought to investigate whether a clearly abnormal response could be identified using either of the gel formulations in pain-free healthy subjects. Finally, Gels A and B were compared in order to determine which was better able to differentiate between normal and abnormal responses.
4.3 **Hypotheses**

1. There will be significant dose-dependent differences in intensity and quality of sensation experienced during the 15-minute application of gel formulations containing different concentrations of menthol.

2. There will be significant dose-dependent differences in temporal characteristics between menthol concentrations: time to onset of first VAS rating, time to maximum rating and overall time-course of sensation intensity and quality.

3. There will be associations between equivalent concentrations of liquid (Chapter 3) and gel (current study) formulations for intensity and quality of response.

4. An abnormal response to menthol will be clearly differentiated by the gel formulations.

5. Construct validity will be demonstrated by good associations between intensity and quality responses to each gel.

6. One of the gel formulations will be more effective in differentiating an abnormal response.
4.4 Methods

Subjects
Twenty-seven pain-free healthy adult subjects were voluntarily recruited for this study from the immediate surrounds of Curtin University, using posters and word of mouth. Inclusion and exclusion criteria were as in Table 4.1. This sample size mirrored the sample used in the previous liquid formulation study.

| Table 4.1: Inclusion / Exclusion criteria for liquid formulation study |
|---------------------------|---------------------------|
| **Inclusion:**            | Aged 18-70                |
|                           | Able to read and understand English |
| **Exclusion:**            | Current pain >1/10 VAS     |
|                           | Currently taking any analgesic or anti-inflammatory medication |
|                           | History of systemic inflammatory conditions; |
|                           | Neurological deficits (motor, cognitive or sensory); |
|                           | History of other chronic pain disorders (eg fibromyalgia, chronic low back pain); |
|                           | Skin allergies;            |
|                           | Allergy to menthol or ethanol. |

All participants provided written informed consent before participating in the study. Ethical approval was provided by Curtin University Human Research Ethics Committee (Approval number PT067/2009).

Study design and procedures
This study followed the same design and procedures as for Study 1 (Chapter 3). A within-subjects, blinded, randomized repeated measures design to evaluate the sensory response of a cohort of healthy pain-free subjects to the two concentrations of menthol gel formulation (Gel A and Gel B). Subjects were blind to menthol concentration.

Subjects attended for testing on 2 occasions, separated by at least 24 hours. All testing occurred in a laboratory with the ambient temperature maintained at 23-24°C. All menthol testing was performed at the same 2x3cm mid-line volar forearm site, 5cms from the wrist crease. The dominant arm was used for all subjects. Standard arm positioning was used, with the subjects maintaining their forearm in full supination.
resting on a pillow placed on a table. Before each test, the site was gently cleaned with tepid water and dried with a paper towel. Before menthol testing, subjects were given 5 minutes to familiarise themselves with the MPQ Descriptor list.

**Preparation of menthol gels:**
All menthol formulation preparation was again carried out under Pharmacy Lab conditions, following standardised procedures. The same batch of solvent was used for both the current study and previous Study 1 (Chapter 3). Menthol solutions A and B were prepared using the same concentrations of menthol as described briefly in Chapter 3. A specific quantity of gelling agent was then incorporated into each solution, using a magnetic stirring device. The gel was left to stir for approximately 1.5 hours at medium speed and then placed in the fridge overnight to reduce air bubbles. Although the gels were stored in the fridge they were brought to room temperature before testing. Gel A contained the same (lower) concentration of menthol as for Liquid A in Study 1; Gel B contained the same (higher) concentration as for Liquid B in Chapter 3.

**Outcome Measures**
Both gels were tested at the same volar forearm site in pre-randomised order on two separate test sessions. A very similar measurement method was used to assess response to the gel formulations as was used for the liquid formulations in Chapter 3, using VAS ratings for intensity and McGill descriptors list for sensation quality. Slight changes were made to assessment procedure based on Chapter 3 findings.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable measured</th>
</tr>
</thead>
</table>
| Response to menthol gel formulations A and B | • *Intensity of cold, heat, unpleasantness and pain (VAS /100)*  
• *Quality of sensation: Pain Rating Index (PRI) and Number of Words Chosen (NWC)* |

*Menthol test application procedure:* Once the test area on the forearm was marked and gently cleaned with hypoallergenic soap and tepid water, 2ml of menthol gel was applied to the site using a 5ml syringe. A Tegaderm dressing was immediately placed over the gel, and the gel then gently spread through the dressing so that it filled the 2x3cm window. A timer was then started for the 15-minute application.

*Intensity ratings:* At one-minute intervals through the 15-minute application, subjects
were asked to rate the intensity of the sensation they were experiencing. An additional VAS scale was added to this study so that intensity of cold, heat, unpleasantness and pain were rated using the same 100mm VAS scales as in Study 1, each line having end-points marked "maximum (sensation) imaginable" and "minimum (sensation) imaginable". A single sheet of paper was used for all four scales, a new sheet provided each minute. Each VAS rating was measured and the value recorded for each sensation at each minute. A final VAS recording was taken at 20 minutes, five minutes after menthol removal.

Quality ratings At two-minute intervals through the 15-minute application, subjects were asked to rate the quality of the sensation they were experiencing. As for Study 1, subjects were asked to select words (no minimum or maximum) from the McGill Pain Questionnaire descriptor list, which the investigator recorded. MPQ Pain Rating Index and Number of Words Chosen index were calculated, with a higher score in both denoting a more intense response. A final descriptor recording was made at 20 minutes.

After 15 minutes, the dressing was removed and the menthol gel immediately washed from the skin using hypoallergenic fragrance-free soap and tepid water, gently patted dry with paper towels. A further recording of VAS intensities and descriptors was taken at 20 minutes.

Data Management and Analysis
Data from intensity and quality rating were managed as follows:

Intensity: VAS values (0-100mm) for intensity of cold, heat, unpleasantness and pain every minute for the 15-minute application were each quantified as:

- Area under the VAS-time curve (AUC) for 1-15 minutes, as a measure of total intensity;
- Maximum VAS value, as a measure of the greatest intensity experienced;
- Time to onset - the time at which the first VAS value >0/100 was recorded;
- Time to peak VAS - the time at which the maximum VAS value was recorded.

Quality: MPQ descriptor list choices for quality of sensation experienced every minute for 15 minutes were quantified using two standard MPQ indices:
• Pain Rating Index (PRI): sum of MPQ ranking values for each different word selected, no single word included more than once.
• Number of Words Chosen (NWC): the total number of different words selected by an individual during the 15-minute application.

Normality testing was carried out using a combination of Shapiro-Wilk tests and visual analysis of distribution. PRI and cold VAS data were normally distributed (Shapiro-Wilk p=.100 to .089) and so parametric analysis was applied. Heat, unpleasantness and pain VAS data was not normally distributed and so non parametric statistics were applied.

Data was then analysed using the SPSS statistical package, version 19. Alpha was set at p<?.05 and the following analyses were carried out:

<table>
<thead>
<tr>
<th>Hypotheses and Related Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Difference in intensity and quality response between menthol concentrations:</td>
<td></td>
</tr>
<tr>
<td>• Area under VAS-time curve for cold, heat, unpleasantness and pain</td>
<td>Paired t-tests or Wilcoxon signed ranks tests</td>
</tr>
<tr>
<td>• Maximum VAS rating for cold, heat, unpleasantness and pain</td>
<td>Descriptive analysis</td>
</tr>
<tr>
<td>• PRI and NWC quality ratings</td>
<td></td>
</tr>
<tr>
<td>• Specific word choice comparison</td>
<td>Pearson’s or Spearman’s</td>
</tr>
<tr>
<td>• Correlations between concentrations</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td>2. Difference in temporal characteristics between menthol concentrations: release times &amp; time-course:</td>
<td></td>
</tr>
<tr>
<td>• Time to 1st sensation (mins) for cold, heat, unpleasantness and pain</td>
<td>Paired t-tests or Wilcoxon signed ranks tests</td>
</tr>
<tr>
<td>• Time to peak sensation (mins) for cold, heat, unpleasantness and pain</td>
<td></td>
</tr>
<tr>
<td>• Time-course for each VAS sensation and key quality descriptors</td>
<td>Descriptive analysis</td>
</tr>
<tr>
<td>3. Association between intensity and quality responses between liquid (Chapter 3) and gel</td>
<td></td>
</tr>
<tr>
<td>(current study) formulations:</td>
<td>Descriptive analysis</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>• Maximum VAS cold, unpleasantness, pain</td>
<td></td>
</tr>
<tr>
<td>• Time-course</td>
<td></td>
</tr>
<tr>
<td>• PRI and NWC</td>
<td></td>
</tr>
</tbody>
</table>

4. Abnormal response determined:  
   • VAS ratings  
   • PRI  
   • Individual descriptors  
   | K-mean Cluster Analysis* |

5. Associations between high intensity and high quality responses to Gels A and B  
   • VAS ratings  
   • PRI  
   • Individual key descriptors  
   | Independent t-tests |

6. One gel formulation will be more effective in differentiating an abnormal response

*Dichotomous cluster analysis was applied to determine normal and abnormal response groups in preference to using a more conventional mean plus two standard deviations cut-off. The rationale for this choice was that inter-subject standard deviations can be large for QST, meaning that a mean+2SD is too high to identify any subject.
4.5 Results

Cohort Characteristics

Twenty-seven subjects participated in this study, 10 males and 17 females. The mean age of the group was 34 years (range 18-52 years). All measures were assessed for gender difference. No significant differences were found for VAS heat, unpleasantness or pain for either Gel A or Gel B (p = 0.517 to p = 0.980) or for release times for either gel (time to onset p = 0.154 to p = 0.989; time to peak p = 0.259 to p = 0.928) There were also no significant gender differences for quality ratings for Gel A (PRI p = 0.425, NWC p = 0.110) or for Gel B (PRI p = 0.994, NWC p = 0.973). There was however a significant group difference in cold ratings for both gels, with males rating maximum cold significantly higher than females: Gel A t_{(26)} = 2.92, p = 0.007; Gel B t_{(26)} = 2.05, p = 0.048.

Hypothesis 1

There will be significant dose-dependent differences in intensity and quality of sensation experienced during the 15-minute application of liquid formulations containing different concentrations of menthol.

- Intensity: there was a significant concentration-dependent difference in VAS intensity as measured by area under the VAS-time curve and by maximum VAS for cold, heat, unpleasantness and pain.

There was a significant difference between concentrations in total VAS intensity (AUC) for each sensation: cold: t_{(26)} = -3.0, p = 0.006; heat: W_{(26)} = -3.25, p = 0.001; unpleasantness: W_{(26)} = -3.17, p = 0.002; pain: W_{(26)} = -3.30, p = 0.001. There was also a significant difference between concentrations for maximum VAS intensity reported during the 15-minute application time (Figure 4.1): cold: t_{(26)} = -2.65, p = 0.013; heat: W_{(26)} = -2.81, p = 0.004; unpleasantness: W_{(26)} = -2.96, p = 0.003; pain: W_{(26)} = -3.07, p = 0.002.

![Figure 4.1: Differences in maximum VAS intensity ratings between Gel A and Gel B.](image-url)
• **Intensity:** There was a concentration-dependent difference in percentage of individuals reporting unpleasantness, heat and pain, but not cold. Gel B resulted in more reports of heat, unpleasantness and pain than Gel A. In contrast, cold was reported by similar percentages of subjects for both concentrations (Figure 4.2).

![Figure 4.2: Percentage of subjects reporting at least some of each VAS sensation (VAS > 0) plus average maximum VAS intensity reported for each sensation at each concentration.](image)

• **Intensity:** there was good intra-subject consistency between menthol concentrations for each VAS sensation

There were moderate to high significant correlations for maximum VAS ratings between concentrations: cold r = .677, p<.001; heat r = .681, p<.001; unpleasantness r = .816, p<.001; pain r = .787, p<.001 (Figure 4.3).

![Figure 4.3: Overlaid scatterplots illustrating positive associations between Gels A and B for cold, heat, unpleasantness and pain ratings.](image)

• **Quality:** There was a significant concentration-dependent difference in word choice, as measured with the PRI and NWC indices.
There was also a significant difference in descriptor choice between concentrations, with the higher concentration provoking higher scores for PRI index and NWC: PRI $t_{(26)}$ = -3.53, $p$ = .002; NWC $t_{(26)}$ = -3.00, $p$ = .006 (Figure 4.4).

![Figure 4.4: Mean quality descriptor index scores for each menthol concentration](image)

- **Quality:** A higher percentage of subjects selected more intense words for Gel B

A dose-dependent effect was found for word choice (Figure 4.5). A higher percentage of subjects selected certain key words during Gel B, rather than different types of words being selected at each concentration. Although cold was selected with similar frequency at both concentrations, the more intense words icy/freezing were chosen by almost twice as many subjects for Gel B and cool chosen more frequently for Gel A. Burning and stinging were also chosen more frequently for the higher concentration Gel B.

![Figure 4.5: Percentage of subjects choosing the most commonly selected words for each gel](image)

- **Quality:** There was good intra-subject consistency between concentrations for PRI and NWC descriptor indices

There was a good correlation between PRI scores for Gel A and Gel B ($r$ = .743, $p$ < .001) and also for NWC score ($r$ = .590, $p$ = .001) (Figure 4.6).
Hypothesis 2

There will be significant dose-dependent differences in temporal characteristics between menthol concentrations: time to onset of first VAS rating, time to maximum rating and overall time-course of sensation intensity and quality.

- There was a significant dose-dependent difference in time to first VAS rating for cold, unpleasantness and pain

Time to onset of first sensation (VAS >0) was significantly quicker for the higher concentration of menthol (Gel B) for cold ($t_{26}=2.15$, $p=.041$), unpleasantness ($W_{26}=-2.43$, $p=.015$) and pain ($W_{26}=-2.49$, $p=.013$). Heat VAS followed the same pattern but the difference did not reach significance ($W_{26}=-1.08$, $p=.219$) (Figure 4.7).

- There was a significant dose-dependent difference in time to first VAS rating for each sensation

Time to peak sensation was reached significantly earlier for the lower concentration gel (Gel A) for each VAS sensation: cold $t_{26}=-2.43$, $p=.022$; heat $W_{26}=-2.49$, $p=.013$; unpleasantness $W_{26}=-2.34$, $p=.019$; pain $W_{26}=-3.08$, $p=.002$ (Figure 4.8).
A similar pattern was shown between concentrations for cold, heat, unpleasantness and pain.

There was a dose-dependent pattern to the time-course of each VAS sensation. Cold VAS showed the sharpest peak and decline for both gels (Figure 4.9a). Unpleasantness and pain followed similar but less sharp rises. Heat VAS followed a slightly different pattern with a lower VAS reported but at a more constant level (Figure 4.9d).

Figure 4.8: Time to maximum VAS for each sensation and each concentration.

Figure 4.9a-d: Mean VAS intensity at each minute during test time for each sensation at each concentration: a) cold VAS; b) heat VAS; c) unpleasantness VAS; d) pain.
• The dose-dependent time-course pattern for key words was less markedly different between high and low concentrations for the gel formulation

For Gel B, cool, cold and icy showed a more consistent pattern of initial rise to peak and decline, while for Gel A these qualities were more inconsistent in their pattern. Burning and stinging showed similar patterns at dose-dependent frequencies.

**Figure 4.10a:** Time-course for selection of key words during Gel B: percentage of subjects selecting each word every 2 minutes.

**Figure 4.10b:** Time-course for selection of key words during Gel A.

**Hypothesis 3**

*There will be associations between equivalent concentrations of liquid (Chapter 3) and gel (current study) formulations for intensity and quality of response.*

• There was little difference in VAS ratings for cold, unpleasantness or pain between Liquid and Gel A or between Liquid and Gel B.

There was minimal difference between formulations in the maximum VAS recorded for each sensation at the 2 different concentrations (Figure 4.11a). Heat sensation was not recorded for liquid formulations and so could not be compared.
There was little difference in sensation quality between Liquid and Gel A or between Liquid and Gel B.

There was also little difference in PRI descriptor choice score between liquid and gel formulations at concentration A or B (Figure 4.11b). Choice of specific words was similar (Figure 4.12), although Liquid B evoked the dysesthetic word tingling and stinging as well as spreading more than Gel B, possibly indicative of the pooling problem reported in Chapter 3.

**Figure 4.11a:** Maximum VAS ratings for cold, unpleasantness and pain for equivalent Liquid and Gel formulations

**Figure 4.11b:** PRI scores for equivalent Liquid and Gel formulations

**Figure 4.12:** Percentage choice of key words for equivalent B concentration Liquid and Gel.

- **Gel formulations at both concentrations tended to show earlier onset and earlier time to peak.**

Although time to onset of cold was very similar between Gels and Liquids, unpleasantness and pain was reported slightly earlier for both gels (Figure 4.13a). The order of onset (cold, followed by unpleasantness the pain) was very similar for all gel and liquid concentrations. Time to peak was clearly earlier for both gel formulations (Figure 4.13b). This may be related to the more even application of the gel formulation.
Hypothesis 4

An abnormal response to menthol will be clearly differentiated by the gel formulations.

- **Intensity**: Dichotomous cluster analysis identified distinct high response groups

Dichotomous K-means cluster analysis for each VAS sensation resulted in a clearly defined “high cluster” group of between 29.76% for cold VAS and 11.1% for pain VAS (Table 4.2). Comparison between mean (SEM) VAS values for high and low cluster groups for each sensation showed that there was a clear distinction between groups for all VAS ratings for both Gel A and Gel B (Figure 4.14a-b).

<table>
<thead>
<tr>
<th></th>
<th>High VAS cluster</th>
<th>Low VAS cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS Cold</td>
<td>8 (29.6%)</td>
<td>19 (70.4%)</td>
</tr>
<tr>
<td>Heat</td>
<td>6 (22.2%)</td>
<td>21 (77.8%)</td>
</tr>
<tr>
<td>Unpleasantness</td>
<td>4 (14.8%)</td>
<td>23 (85.2%)</td>
</tr>
<tr>
<td>Pain</td>
<td>3 (11.1%)</td>
<td>24 (88.9%)</td>
</tr>
</tbody>
</table>
**Intensity:** when individual VAS cluster group membership was compared, 6 subjects (22.2%) showed a widespread atypically intense response.

Analysis of individuals who were classified in high cluster groups showed that 3 subjects (11%) were classified as atypical for all 4 VAS sensations. A further three were classified as high cluster or close to high cluster for unpleasantness or pain plus either heat or cold (Table 4.3).

**Table 4.3:** Individual membership of high cluster groups for each VAS sensation at each concentration of gel (✓). Mean VAS given when individuals did not meet the VAS cut-off value for high cluster membership.

<table>
<thead>
<tr>
<th>Subject</th>
<th>A</th>
<th>B</th>
<th>A</th>
<th>B</th>
<th>A</th>
<th>B</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>23</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>21</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>25</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>27</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>10</td>
<td>35</td>
<td>30</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Quality:** Dichotomous cluster analysis identified distinct high response groups

As shown in Table 4.4, the high PRI cluster comprised 8 subjects (29.6%), with a cut-off score of 9 for Gel A and 10.5 for gel B. There was a large difference in mean PRI scores between cluster groups (p<.001) showing their distinctly different characteristics (Table 4.4).
Table 4.4: Number (percentage) of subjects classified in high or low PRI cluster groups following dichotomous K-means cluster analysis

<table>
<thead>
<tr>
<th></th>
<th>High PRI cluster group</th>
<th>Low PRI Cluster group</th>
<th>PRI cut-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
<td>8 (29.6%)</td>
<td>19 (70.4%)</td>
<td></td>
</tr>
<tr>
<td>Gel A</td>
<td>12.0 (1.3)</td>
<td>3.8 (0.54)</td>
<td>9</td>
</tr>
<tr>
<td>Gel B</td>
<td>16.0 (1.8)</td>
<td>5.9 (0.55)</td>
<td>10.5</td>
</tr>
</tbody>
</table>

When individual word choices of subjects in each PRI cluster group were explored, high PRI groups for both concentrations consistently selected a limited number of words (Figure 4.15). The most discriminatory words were icy, burning, prickling and stinging.

![Figure 4.15: Percentage of subjects in high and low cluster groups selecting key words.](image)

**Hypothesis 5**

**Construct validity will be demonstrated by good associations between intensity and quality responses to each gel.**

- Independent t-tests showed that the high PRI cluster group showed significantly higher VAS ratings for unpleasantness and pain for both gel concentrations.

There were significant differences between high and low PRI groups for unpleasantness and pain at both gel concentrations: Gel A unpleasantness $t_{(26)} = 2.21$, $p=.036$; pain $t_{(26)} = 3.59$, $p=.001$; Gel B unpleasantness $t_{(26)} = 2.51$, $p=.019$; pain $t_{(26)} = 4.39$, $p<.001$. There were no significant cluster group differences for cold (Gel A $t_{(26)} = 1.06$, $p=.298$; Gel B $t_{(26)} = 1.35$, $p=.189$) or heat (Gel A $t_{(26)} = .407$, $p=.688$; Gel B $t_{(26)} = .7689$, $p=.497$).

- Subjects classified as high intensity for VAS and PRI were more likely to select the words icy, burning and stinging.
Although the subject numbers are small, those who show high scores for both intensity and quality were more likely to choose the key words icy, burning and stinging, particularly for Gel B. Those who were members of the high cold VAS cluster only chose the intense cold words (Table 4.5).

Table 4.5: Classification of subjects as cold hyperalgesic or not according to membership of high VAS cluster groups, high PRI cluster groups and choice of key words. VAS / PRI score > than the cut-off is designated ✓ VAS values < the cut-off are recorded.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Member of VAS high cluster</th>
<th>Member of PRI high cluster</th>
<th>Word choice</th>
<th>Cold Hyperalgesic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cold</td>
<td>Heat</td>
<td>Unpl</td>
<td>Pain</td>
</tr>
<tr>
<td>14</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>23</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>21</td>
<td>✓</td>
<td>0</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>25</td>
<td>✓</td>
<td>0</td>
<td>✓</td>
<td>15</td>
</tr>
<tr>
<td>27</td>
<td>✓</td>
<td>✓</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>✓</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>0</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>25</td>
<td>0</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>10</td>
<td>✓</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>✓</td>
<td>0</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>✓</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Hypothesis 6**

**One of the gel formulations will be more effective in differentiating an abnormal response**

- Although both gels identified abnormal responses, the percentage of subjects in Gel B high cluster groups for VAS and PRI were more consistently between 15% and 25% of the total group.

K-means cluster analyses were performed for each gel for each VAS sensation and PRI value. For Gel B, there were between 15 and 22% of subjects in the high cluster groups for PRI and for VAS heat and unpleasantness. For Gel A the percentage of subjects in the high cluster groups for each VAS sensation and PRI was either below or above the target 15-25% zone (Figure 4.16).
Figure 4.16: Percentage of subjects in the high score cluster for each concentration.
Horizontal lines designate a proposed target zone for a cold hyperalgesic group between 15 and 25% of healthy subjects.

- Twenty-six percent of subjects were members of all five high cluster groups. Gel B was better able to identify these subjects.

Individual membership of each high cluster group was compared. For Gel B, 3 subjects fulfilled membership criteria for all 5 variables (subjects 1, 14 and 23 in Table 4.6). These subjects also selected at least 2 / 3 key descriptors icy/freezing, burning and stinging. 4 other subjects were members of at least 4 / 5 of the variables and selected at least 2/ 3 key descriptors (subjects 6, 9, 19 and 21). For Gel A, no subject fulfilled membership of all 5 variables.

Table 4.6: Membership of each VAS and PRI high cluster group (with cut-off values shown) and identification of choice of key words icy, burning and stinging. VAS values shown when cut-off value not reached.

<table>
<thead>
<tr>
<th>Sbjt</th>
<th>VAS criteria</th>
<th>PRI score 15+</th>
<th>Word choice includes:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cold 30+</td>
<td>Heat 20+</td>
<td>Unpl 20+</td>
</tr>
<tr>
<td>Gel B</td>
<td>14</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>✓</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>✓</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>20</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0</td>
<td>✓</td>
</tr>
<tr>
<td>Gel A</td>
<td>14</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>✓</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>✓</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>
### Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Accepted/Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. There will be significant dose-dependent differences in intensity and quality of sensation.</td>
<td>Accepted</td>
</tr>
<tr>
<td>2. There will be significant dose-dependent differences in temporal characteristics.</td>
<td>Accepted</td>
</tr>
<tr>
<td>3. There will be associations between equivalent concentrations of liquid (Chapter 3) and gel (current study) formulations for intensity and quality of response.</td>
<td>Accepted</td>
</tr>
<tr>
<td>4. An abnormal response to menthol will be clearly differentiated by the gel formulations.</td>
<td>Accepted</td>
</tr>
<tr>
<td>5. Construct validity will be demonstrated by good associations between intensity and quality responses to each gel.</td>
<td>Accepted</td>
</tr>
<tr>
<td>6. One of the gel formulations will be identified as more effective in differentiating an abnormal response</td>
<td>Accepted</td>
</tr>
</tbody>
</table>
4.6 Discussion

This study evaluated in normal subjects the dose-dependent effects of two concentrations of a menthol gel formulation, developed as a result of findings from Chapter 3 of this investigation. Intensity and quality responses were compared descriptively with those from equivalent liquid formulations from Chapter 3. Internal construct validity for the two gels was also assessed. Finally cluster analysis was applied to determine whether an abnormal response could be clearly differentiated in this healthy population.

Dose-dependent effects of menthol gel

This study of two gels with lower and higher concentrations of menthol showed good dose-dependent effects, in a similar manner as was found for the liquid formulations. Menthol Gel B evoked significantly higher VAS intensity for all sensations. Analysis of individual word selection showed that similar types of words were chosen at both concentrations but more subjects chose them for Gel B. For example: icy/freezing was chosen by 48% of subjects for Gel B but by only 26% for Gel A; burning, stinging and prickling were also more frequently selected during Gel B application (Figure 4.5). Sensations associated with the higher concentration Gel B was also experienced significantly earlier for cold, unpleasantness, and pain. Although VAS heat showed a similar pattern, the difference between concentrations did not reach significance. This suggests that more menthol was penetrating the skin and activating receptors with Gel B than with Gel A. Moderate to high significant correlations were seen between concentrations for each VAS sensation (Figure 4.3) suggesting good response consistency.

These results suggest that the proposed menthol gel formulation therefore provided a consistent, dose-dependent delivery, and sensory activation, indicating its potential value as a reliable standardised test application.

Comparison with liquid formulations

Since there are no other studies using a menthol gel against which to evaluate these results, comparisons were made with the equivalent liquid concentrations from Study 1 (Chapter 3) in order to assess whether addition of the gelling agent had influenced intensity, quality or release times. The same base solvent was used for both studies to reduce variability. This descriptive comparison showed that there was remarkably little difference between gel and liquid for either intensity or quality response.
Comparison between maximum VAS values (Figure 4.11a) showed remarkable similarity: for example mean cold VAS was 32.2 for liquid B and 35.0 for Gel B; mean unpleasantness VAS was 11.3 for Liquid A and 11.9 for Gel A. Similarly there was little difference in PRI descriptor scores, although the higher concentration Liquid B evoked a slightly higher score than Gel B. Closer analysis of word choice showed that the dysesthetic words tingling, stinging and spreading were used more often for the liquid formulations, in particular for Liquid B. This difference may reflect the pooling problems of the liquid formulation rather than reduced effect of the gel formulation. Anecdotal report of liquid pooling in Chapter 3 was associated with patches of more intense sensation and higher report of spreading and stinging descriptors, particularly when an individual was unable to maintain forearm supination. Gel delivery improved this, resulting in no reports of particularly intense areas of sensation and no visible flow of gel to the dependent edge when an individual moved their forearm into a more neutral position. Aside from the words spreading and stinging, frequency of choice of milder cold-type words and burning was similar (Figure 4.12).

Comparison of release characteristics however revealed some variance between formulations which may reflect the different formulation. On the one hand the time to onset was strikingly similar between gel and liquid indicating that addition of the gelling agent has no slowing effect on initial release. The higher concentration always initiated a response slightly earlier and the order of sensation was identical (Figure 4.13a), with cold always first and recorded at a very similar time-point for both formulations. In all cases unpleasantness was next, although slightly later for liquid formulations and pain was always the last sensation to be felt. There was however a clear and systematic difference between formulations in time to peak. Both Gel A and B reached maximum VAS for all sensations considerably earlier than the equivalent liquid formulations (Figure 4.13b). Although this finding is surprising, it demonstrates that a gelling agent does not slow release of menthol; if anything, it may have enhanced release. Gel formulations therefore appear to offer more consistent and perhaps slightly more rapid release than the liquid formulation, whilst maintaining similar levels of sensation intensity and quality.

**Validity**

Internal content and construct validity was also maintained in the new gel formulation, tested with an entirely different cohort of participants. Both gels evoked a mainly cold
sensation with some additional unpleasant dysaesthetic sensations, similarly to the liquid formulation findings. The time-intensity curve for each sensation followed a similar pattern for each gel concentration (Figures 4.10a and b), which in turn was similar to the pattern for liquid formulations (Figures 3.11a and b). A strikingly similar and limited set of descriptors was selected from the long McGill list for both concentrations at both formulations. These findings all suggest that a similar construct was being tested and reported for both formulations. Internal construct validity was also indicated by the good relationship between VAS intensity scores and word choice. When subjects were divided into low and high PRI groups according to cluster analysis, those scoring higher PRI values also reported significantly higher intensity of unpleasantness and pain than those in the low PRI group. Cold and heat VAS were not significantly different between PRI groups. However, this is an additional sign of the appropriateness of the relationship since the PRI gives low weightings to less intense cold or heat-type words whilst giving a higher weighting to more unpleasant or painful words such as burning, icy or stinging. This pattern of association therefore adds to the internal validity of the menthol test.

Assessing comparative validity against other studies is difficult as these studies are scarce and have been focussed on the sensitising influence that menthol then has on other measures rather than on the effect of menthol per se. All studies also involve very small cohort numbers. Menthol formulations, delivery methods and application times in these studies are variable and reported response to menthol very short on detail. Two studies have used a low concentration of menthol and have reported only low levels of cooling (Eccles, 1994) plus, in some, mildly noxious sensations. Green and Schoen (2007) used a Labeled Magnitude Scale to report ‘weak’ to ‘barely detectable’ cooling and nociceptive sensations, but no additional details are reported. Hatem et al. (2006) used 30% menthol as a sensitising agent for subsequent thermal cold stimulation, but still reported that 90% of subjects spontaneously reported only a cool sensation and 10% a warm sensation after 10 minutes application. Two studies applied high concentration 40% menthol with healthy subjects. Both found that at least 80% of subjects reported pain (mean NRS 2/10 (Binder et al., 2011)); NRS 3/10 (Wasner et al., 2004). Both studies also reported that burning was the main quality associated with noxious sensation, although incidence varied widely: 30% for the Binder study and 88% for Wasner. Similarly Binder et al. (2011) reported only minimal incidence of coldness (17%) whereas Wasner reported in more detail that all subjects reported coldness of mean 4.5 on NRS which started two minutes after application. The more
elevated response results from Wasner et al. (2004) are more in accordance with the findings of the current study.

**Identification of an abnormal response**

This study also investigated whether the current gel formulation of menthol was able to differentiate between a normal and abnormal response in a healthy cohort. As already implied, no comparison with previous data is possible as no study has systematically reported the range of intensity and quality responses for a sufficiently large normal cohort. Two-group cluster analysis was therefore applied to each of the VAS and PRI variables in the current study. This approach has been used in a recent study aiming to quantify with a single cut-off value an abnormal cold response using CPT (Sterling et al., 2011). The current study had the added bonus of being able to use the same cluster analysis approach and apply it to both VAS intensity data and to quality data that would provide information about normal and abnormal descriptor choice.

Each VAS sensation cluster analysis identified a small group of higher scoring subjects: around 30% of the total cohort for high cold VAS, 15-18% for high heat and unpleasantness and 10% for high pain. PRI index clustering also resulted in a high score group of around 30% of subjects. Confidence intervals and SD values for high and low cluster groups in all cases showed minimal overlap between groups, suggesting clear differentiation. However, an abnormal response to cold will involve a combination of these intensity and quality variables. When individual membership of each high cluster group was analysed for overlaps, 3 subjects (11%) were found to be members of the high PRI and at least 3 of the 4 high VAS groups (Table 4.5). A further three subjects were very close to the cluster cut-off values for at least three of these variables. This indicates that 22.2% of the normal cohort showed a number of elements of a significantly abnormal response. This percentage is in accordance with the percentages of normal subjects who have been reported to exhibit abnormal pain responses to cold (Wright et al., 2010)¹.

**Characterisation of an abnormal response**

¹ See Appendix 5
Comparing these six subjects with the remaining 21 more normal responders, normal and abnormal responses to menthol could be characterised. Cold was experienced as the initial and predominant sensation for all but one subject even though the intensity of cold was relatively low: the low cluster group recorded a maximum VAS intensity of only 24 (± 2.2)/100 for Gel B in contrast to 62(±6.1)/100 for those in the high cluster group. This result supports the basic science finding that menthol predominantly activates TRPM8 receptor channels on Aδ thermal fibres which signal cool or cold sensations (Section 2.6). This cold sensation was reported throughout menthol application, even in the presence of additional more noxious sensations. It has been suggested that TRPM8 channels on Aδ thermal fibres are reciprocally deactivated as noxious c-fibre activation increases in the presence of higher intensity cold or menthol (Babes et al., 2010). Yet the findings from this study indicate that cold sensation continues to be experienced concurrently with more noxious sensations. It may be that this is a reflection of activation of the high threshold cold sensitive c2-fibres reported by Campero et al. (2009) whose proposed role is to signal a noxious change as a transition before more intensely noxious polymodal c-fibres start to fire.

Heat was also reported by a surprisingly high percentage of subjects during the application of menthol (30% of subjects during Gel A and 35% during Gel B). However, for the majority of individuals in the low cluster, heat VAS peaked at a negligible 8/100. Those in the high cluster recorded maximum VAS heat values of 29(±6.8)/100 for Gel A and 35(±5.5)/100 for Gel B. A ‘normal’ response to menthol therefore may include report of very low levels of heat. Anecdotally individuals report a non-noxious hot-cold sensation which is difficult to differentiate. This supports a previous study which reports warm sensations during sustained cold temperature (Davis & Pope, 2002).Paradoxical heat sensation (during cold application) has been proposed as a sign of nociceptive activity since TRPV1 heat receptors are only expressed on nociceptors (McKemy, 2012). Paradoxial heat may also be a sign of the disinhibition of Aδ cool fibres which usually inhibit heat transmission at spinal cord level (Fruhstorfer, 1984).Experiments using the thermal grill illusion have proposed that paraxoical heat may also be a sign of more centrally-controlled disinhibition, controlled by changes in spinal cord opioid channel expression (Craig & Bushnell, 1994; Jorum et al., 2003; Kern et al., 2008). In the current study of healthy subjects however, some degree of non-noxious heat accompanying a cold sensation appears to be normal. As intensity increases (Gel A to Gel B) the cold does not reciprocally abate as heat steadily increases but instead for those with a more intense response, there appears to be a change in
quality to a clearly noxious burning sensation. This suggests a different mechanism to
the non-noxious warm-cold, more reflective of nociceptive activity. It may therefore be
that the key abnormal component is the burning sensation rather than heat per se.

Similarly to heat, although high numbers of subjects reported some VAS
unpleasantness and pain at both Gel A and B concentrations, the low cluster group
rated these at negligible levels of intensity (7-12/100 for unpleasantness and 1-5/100
for pain). In contrast, those in the high cluster groups rated unpleasantness at 40-
55/100 and pain at 15-45/100. It therefore seems that, similarly to heat, a normal
response to menthol may include an element of very low intensity unpleasantness or
pain. In clear contrast, those exhibiting an abnormal response will rate intensity at
>40/100 for unpleasantness and >20/100 for pain.

Specific word choice appears to provide an additional and important discriminatory
element. A strong relationship can be seen between abnormally high intensity of
unpleasantness or pain and selection of the noxious words burning, stinging and
icy/freezing (Table 4.5). Although the word burning may imply activation of noxious
heat receptors, in particular via TRPA1 which is active during inflammation (Schaible
et al., 2011; Bautista et al., 2012) this is less likely to be the mechanism in the case of
individuals without pathology. This abnormal response may instead imply an
underlying dysfunction in descending inhibitory pain systems. Several studies have
shown that A-fibre blocks result in a previously cool sensation being experienced in a
very similar manner to the abnormal response of the current study, as an unpleasant
icy, stinging and burning sensation (Frustorfer, 1984). Green et al. (2008) also
describe cold spots in the skin that evoke a noxious stinging sensation when stimulated
with non-noxious cold and propose that this is a sign of dysfunctional central
inhibition. It may therefore be that the abnormal response demonstrated by the 6
subjects in the current study reflect a pre-existing altered inhibitory drive. This has
been reported in pain-free subjects in a number of studies. For example, Yarnitsky et al.
(2008) found that reduced DNIC efficiency predicted poor outcome post thoracotomy
surgery.

This study found that no single measure of intensity or quality provided sufficient
specificity to identify a cold hyperalgesic response on its own. An abnormal response to
a normally non-noxious menthol-cold stimulus involved elevated VAS unpleasantness
and pain ratings associated with a limited set of words to describe the noxious quality.
A measurement outcome that combines both intensity and quality domains is better able to reflect the multi-faceted nature of the hyperalgesic response and has the potential to provide improved specificity and sensitivity. The menthol test for cold hyperalgesia therefore has potentially better content validity than the conventional CPT temperature value. This is particularly the case if the two aspects of sensory response could be combined.

**Which concentration best identified normal and abnormal groups?**

This study compared Gels A and B with the aim of determining the optimal concentration for a test for cold hyperalgesia. The goal of such a test is to be able to identify individuals who have the early signs of centrally-augmented pain and are at risk of developing more firmly established problems, not just those who already have structural pain processing abnormalities. The final test formulation therefore needs to be non-noxious for the majority of subjects but of sufficient intensity to ensure that more than just the very extreme hyperalgesic individuals are identified.

Both menthol gels identified a group of individuals with an elevated response, although with different percentages of abnormal individuals identified for each concentration. Studies that have identified individuals with cold hyperalgesia using a conventional thermode have reported the incidence to be around 15-25% in healthy subjects and anything from 30% and 100% in different pain pathologies (Sterling et al., 2006; de la Llave-Rincon et al., 2009). For the current study, when K-mean cluster analyses were carried out for each menthol concentration, the percentage of subjects in the Gel B high cluster groups were more often within the ideal 15-25% range than for Gel A (Figure 4.16). This suggests that the Gel B concentration was consistently better able to identify the target percentage of hyperalgesic subjects across all measures. When membership of high cluster groups for PRI and VAS variables was compared, 7 subjects (25.9%) fulfilled at least three of the five VAS and PRI characteristics plus selected at least two of the three the key descriptors (Table 4.6). Gel A identified only two of these subjects. Although this percentage is higher than anticipated in a normal cohort, it illustrates that, although Gel A offered some discriminatory ability, Gel B was superior.

A valid test for hyperalgesia needs to be able to identify reliably those individuals who exhibit both high intensity values and high quality index scores. Although both gel concentrations were able to identify individuals with an elevated response, the higher
Concentration Gel B more consistently identified the most atypical group. Gel B will therefore be the test formulation applied in future studies.

**Assessment of test procedure, practical and safety issues:**

In addition to scientific efficacy, this study also evaluated the safety and practicality of the gel menthol test, using anecdotal evidence from both the investigator and from participants.

- **Adverse responses**

There was a low incidence of erythema for either Gel A or B and it did not cause discomfort to any subject. The two subjects with a persistent erythema reported its disappearance 30 minutes post removal. No other adverse reactions were reported.

**Stimulus consistency**

Given the problems with pooling for the liquid formulation, close attention was taken in the current study to sensation consistency across the test area. The majority of participants reported that they were not aware of any particular differences in sensation intensity. It was therefore concluded that adding gelling agent provided a more stable vehicle for the menthol.

- **Test practicalities**

The delivery method was relatively easy to apply, although using a syringe to draw out 2mls of gel was messy and clearly inefficient for use in a clinical setting. A future commercial test prototype will ideally enclose the gel in an adhesive device that can be quickly applied to the skin and immediately release the gel. The paper VAS and modified descriptor lists were manageable for research purposes but, again, inefficient for a clinical environment. In particular the paper VAS scales were unwieldy for participants to record four different sensation intensities: a new page of VAS scales was needed every minute and the scales needed to be measured and added to a spreadsheet manually. An electronic VAS system would improve efficiency.
4.7 Summary

This study added a gelling agent to the basic menthol solution at two concentrations and found that dose-dependent effects for intensity, quality of sensation and release times in healthy subjects remained intact whilst delivery consistency improved. Levels of intensity for cold, unpleasantness and pain and descriptor choices were similar between equivalent gel and liquid concentrations from Study 1 (Chapter 3). Words such as stinging and spreading were selected more frequently for liquid formulations, but this was probably more associated with liquid application issues than with reduced gel effect. Overall, the gel formulations showed good content and construct validity. Release times for the gel formulation were slightly quicker than for the liquid, meaning that the gel would be suitable for clinical application.

Response characteristics for a typical / normal and atypical / abnormal response were analysed. A normal response to a moderate concentration of menthol involves a clear sensation of cool or cold temperature, possibly associated a low intensity of heat or mild unpleasantness and reports of mild dysaesthetic sensations such as tingling. An abnormal response is clearly differentiated by higher intensity unpleasantness and pain (+/- heat and cold) combined with a high PRI descriptor score which will include at least two of the three key words icy/freezing, burning and stinging. This abnormal response was found in approximately 25% of this healthy group and may reflect a mildly dysfunctional descending pain inhibitory system.

Future studies will need to improve the practicalities of the test, for example by developing an electronic VAS system. Although the PRI index has value, it was developed to describe spontaneous pain and concentrates solely on quality descriptors. The current study found that combination of intensity and quality ratings provided a more comprehensive measure of response. Further development of an index that combines weighted descriptor scores with VAS intensity ratings would be valuable. Finally, this new index then needs to be validated in a conventional cold-temperature study before being applied to a population with pain pathology.
Chapter 5

Study Three

Validation of the ADI scoring system using cold thermal stimuli

5.1 Abstract

Background and Aims
The ADI scoring system, which combined measures of sensation intensity and quality, was developed to differentiate between normal and abnormal responses to cold stimuli. Development was based on data from previous menthol studies (Studies 1 and 2). The primary aim of this study was to validate the ADI scoring system in the context of conventional thermal cold stimuli. Firstly the ability of the ADI to discriminate between three sustained cold stimuli (10°C, 15°C and 20°C) was assessed. Criterion validity was also assessed by comparing ADI with McGill Pain Rating Index score at each cold stimulus. Associations between cold pain thresholds values and ADI score were also assessed.

Method
Twenty-nine healthy volunteers (mean age 34 years) were assessed on a single occasion. All testing was performed on the volar surface of their dominant forearm. Initial cold pain threshold assessment was carried out in triplicate using a Medoc TSA II thermode and standard method of limits, with the mean used for analysis. Subjects were also divided into groups according to CPT< or >15°C. Three sustained cold stimuli (10°C, 15°C and 20°C) were then applied for five minutes each, in randomised order. During application, subjects were asked to complete VAS rating scales for intensity of cold, heat, unpleasantness and pain every 30 seconds. Every one minute subjects were also asked to select words from a MPQ descriptor lists to describe the sensation experienced. The ADI was calculated from these values, in addition to a McGill PRI score.

Results
There was a significant difference in ADI between temperatures (p<.001), with higher scores at lower temperatures. The individual ADI components also showed significant temperature-dependent effects. Word choice was clearly different: at 10°C more subjects reported cold, icy/freezing, burning and prickling; at 20°C subjects reported more cool and warm sensations. ROC curve analysis showed that ADI was a more effective predictor of CPT group (<15°C) than PRI: AUC .832, sensitivity .90, specificity
.74 at a cut-off score of 4.9. Those with CPT>15°C exhibited significantly higher ADI scores, reporting higher MWS descriptor scores and greater intensity of heat, unpleasantness and pain at 10°C and 15°C, although no difference in cold VAS. There were strong similarities between sustained 10°C response and response to Gel B.

**Conclusion**

The ADI showed good ability to discriminate between sustained thermal cold temperatures and good internal construct validity. The similarities in ADI-measured response between sustained cold and menthol gel also supports the use of topical menthol test as an alternative sustained cold stimulus. ADI was also able to predict membership of a cold hyperalgesic group (defined with CPT>15°C), suggesting that ADI is able to discriminate between normal and abnormal responses using a cut-off of 4.9.
5.2 Introduction & Background

Studies 1 and 2 demonstrated that the response of healthy individuals to topical menthol varies according to concentration and there are clearly identifiable ‘normal’ characteristics to that response in terms of intensity and quality of sensation. It therefore appears possible to identify an atypical response.

However, accurate differentiation between normal and abnormal cold responses is dependent on a suitable stimulus matched with a sensitive measurement tool. Tests for cold hyperalgesia have been previously developed, some quite recently. However, these tests either lack good reliability, for example with the thermode or with use of a cool object (Baron et al., 2010; Eaton et al., 2012; Moloney et al., 2012) or suffer from problems with validity, for example with a test involving application of a supra-threshold ice stimulus (Maxwell & Sterling, 2012). In the current investigation, Study 2 determined that a specific gel formulation (Gel B) provided an alternative cold stimulus that was stable and contained a concentration of menthol consistently non-noxious to the majority of healthy individuals, but that also provided a sufficiently intense stimulus as to evoke clearly identifiable noxious characteristics in a small minority. The size of this atypical group was consistent with the proportion of abnormal cold responses that might be anticipated in normal subjects (Wright et al., 2010)1. Any stimulus however is reliant on its measurement device. This needs to be sufficiently sensitive as to be able to differentiate between typical and atypical responses and sufficiently specific as to be able to accurately identify the particular phenomenon being assessed. Several of the recent studies applying different cold stimuli have lacked well-developed measurement approaches and so found problems in clearly identifying the complex phenomenon of cold hyperalgesia (Hatem et al., 2006; Eaton et al., 2012; Maxwell & Sterling, 2012).

Although cold pain threshold is still the ‘gold standard’ approach to identifying a hyperalgesic response to cold, its viability is hampered by its measurement approach. CPT relies entirely on a single averaged temperature value as its measure of cold pain response. Accuracy and reliability is problematic because CPT generates a uni-dimensional temperature value to reflect a multi-dimensional phenomenon. It is assumed that CPT value reflects a specific response (noxious cold) and the temperature selected by the subject is indicative of whether that noxious response is appreciated or

1 Refer Appendix 5
not. However in practice it is difficult to know how accurate this assumption is because no additional information is sought from the individual.

A measurement approach that combines information from more than one pain domain may potentially offer better content validity, particularly in the context of assessing central pain augmentation. In Studies 1 and 2 of the current investigation a rudimentary measurement system was developed that aimed to combine the domains of intensity and quality of sensation experienced during the menthol stimulus. VAS scales for cold and also heat, unpleasantness and pain were included. It was anticipated that a typical response to a higher concentration of cold stimulus would involve increased intensity of sensation, a ‘turning up of the volume’. An atypical response, for example from an individual with augmented pain processing, would be seen as a far greater increase in intensity (hyperalgesia). This is reflected in basic science. Reduction in cold temperature or increases in concentration of menthol results in greater influx of calcium into TRPM8-expressing neurons (Mckemy, 2007). This is experienced as increased cold intensity. Response to cold may also involve an element of warmth (Harrison & Davis, 1999) and as the response becomes more atypical unpleasantness and pain sensations are triggered at higher levels of intensity.

However, as cold stimulus intensifies still further, a normal response changes in quality of sensation as well as in level of intensity (Davis & Pope, 2002; Binder et al., 2011). This change in mode is important mechanistically as it signals that a different pathway has been triggered. For example, very intense cold may be experienced as a numbing ache rather than cold, or as a combined burning-freezing sensation. This may be described as unpleasant or painful but the key factor is that the quality has changed from an expected mode (cold) to one that is associated with a different sensory pathway (burning). In individuals with an abnormally sensitised or augmented response, change in quality of sensation at a normally non-noxious cold temperature is likely to be an important indicator that the sensory signal system is dysfunctional. This may be the result of multiple factors including heterosynaptic potentiation or dysfunctional spinal inhibitory systems or of phenotypic changes in the dorsal horn. Although measurement of sensory quality is not able to differentiate the specific dysfunctional mechanism, it does signal centrally-mediated dysfunction in a way that intensity alone cannot.
Studies 1 and 2 used both the McGill descriptors list and VAS scales to quantify response to the menthol stimulus. Although these measures were successful in differentiating between concentrations, individually each was less successful in differentiating between individuals with normal or abnormal responses. Combining the domains of intensity and quality appeared to provide a more sensitive and specific measure. Consequently, the current investigation developed a new index, known as the Algotect Descriptor Index (ADI), which aimed to refine the idea by combining VAS scales for intensity of cold, heat, unpleasantness and pain with key elements of the descriptor list. VAS scales were weighted according to anticipated value in identification of an abnormal response to a cold stimulus. Individual words were weighted using a similar theoretical approach, informed by previous pilot work and results from Studies 1 and 2. Appendix 2 outlines the basic approach taken, although the specific details of weightings and method of calculation of the final index score is not included, for reasons of protecting intellectual property.

The ADI scoring system therefore involves rating of the four VAS intensities and selection of words from a descriptors list at regular intervals during a cold stimulus. A total ADI index score is calculated by combining weightings for maximum VAS for each intensity with the sub-score for the descriptor element (maximum score 9). This scoring system however needs to be validated using known cold stimuli. Conventional thermal-cold stimuli at different temperatures are accepted as evoking a graded cold sensation. Therefore if the ADI is a valid measure of cold response, it should be able to discriminate between graded cold temperatures with clearly temperature-dependent scores. It would be anticipated that differences would be shown in the intensity element of the ADI and also in the descriptor element. Sustained temperature stimuli were chosen partly so that individuals had sufficient time to compute accurately the sensation experienced, but also so that comparisons could be made with sustained menthol stimuli. The temperatures 20°C, 15°C and 10°C were selected for a number of reasons. The human cold sensory system is able to differentiate between hot temperatures of less than one degree, largely due to the existence of a range of graded heat channels. Cold temperature discrimination is less precise, particularly as the temperature reduces below 25°C (McKemy, 2012), reflecting the greater complexity of the cold signaling system. Therefore the three temperatures were separated by 5°C intervals. Twenty degrees centigrade is normally felt as non-noxious cool, whereas 10°C is at the upper limit of normal cold pain threshold (Bennett, 2006). Fifteen degrees centigrade has been proposed as a potential cut-off temperature for activation
of more noxious pathways, possibly involving TRPA1-expressing c-fibres (Campero & Bostock, 2010).

The primary aim of this study therefore was to validate the ADI scoring system in the context of conventional thermal cold stimuli. This involved investigating whether the ADI scoring system was able to discriminate between three sustained cold temperatures: 10°C, 15°C and 20°C. In addition, criterion validity was assessed by comparing ADI scores with McGill PRI scores and investigating associations between cold pain threshold values and ADI scores.

5.3 Hypotheses

1. The ADI scoring system will discriminate between three sustained cold temperatures.

2. The ADI scoring system will show greater accuracy than the McGill PRI index in discriminating between temperatures

3. There will be significant associations between cold pain threshold value and ADI score at each sustained cold temperature:
   a. There will be positive correlations between CPT and ADI
   b. Those with high CPT will show significantly higher ADI scores for each sustained temperature

4. There will be similarities in intensity and quality of response to 10°C sustained cold and menthol Gel B (from Study 2).

5. There will be no association between psychological measures and either CPT value or ADI at any sustained temperature.
5.4 Methods

Subjects
Twenty-nine pain-free healthy adult subjects were voluntarily recruited for this study, using posters and word of mouth. There were no previous studies comparing response to sustained cold at different temperatures with which to calculate a sample size. Previous studies applying a single sustained cold temperature to different skin types and comparing responses from the MPQ descriptor list (Davis, 1999; Davis & Pope, 2002) used sample sizes of 21 and 17 respectively. A sample size of approximately 30 was therefore considered sufficient for the range of data needed for the current study. Inclusion and exclusion criteria were identical to those for Studies 1 and 2 and are shown in Table 5.1.

| Inclusion: | Aged 18-70
|           | Able to read and understand English |
| Exclusion: | Current pain >1/10 VAS
|           | Currently taking any analgesic or anti-inflammatory medication
|           | History of systemic inflammatory conditions;
|           | Neurological deficits (motor, cognitive or sensory);
|           | History of other chronic pain disorders (e.g. fibromyalgia, chronic low back pain);
|           | Skin allergies;
|           | Allergy to menthol or ethanol. |

All participants provided written informed consent before participating in the study. Ethical approval was provided by Curtin University Human Research Ethics Committee (Approval number PT106/2007).

Study design and procedures
This study used a cross-sectional observational design to evaluate the sensory response of a cohort of healthy pain-free subjects to three sustained cold temperature stimuli using the ADI scoring system.

Subjects attended a single test session. Psychological questionnaires were completed first. The test procedure was explained in detail to subjects who were then given five-minutes to familiarise themselves with the descriptor word list and the electronic VAS
device. Test order was partially standardised, with threshold testing always carried out before sustained temperatures. Cold detection threshold (CDT) was always assessed before cold pain threshold (CPT) to ensure intact neural pathways. Following removal of the thermode probe and a five-minute pause to allow skin sensation to return to normal, three sustained cold temperatures (10°C, 15°C, 20°C) were applied in randomised order. Each sustained temperature was applied for a total of five minutes, with a five-minute rest at baseline temperature between stimuli. Subjects were not informed about the exact temperature of each sustained stimulus.

**Outcome Measures**

All outcome measures were tested at the same volar forearm site. Subjects were seated at a table, with their dominant arm resting in supination with the volar surface uppermost. The forearm test site was first marked at 5cms from the wrist crease in mid-line.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response to brief cold</td>
<td>• Cold pain threshold (CPT) (°C)</td>
</tr>
<tr>
<td>Response to sustained cold temperatures 10°C, 15°C, 20°C</td>
<td>• Intensity of cold, heat, unpleasantness and pain (VAS /100)</td>
</tr>
<tr>
<td></td>
<td>• Quality of sensation:</td>
</tr>
<tr>
<td></td>
<td>o Pain Rating Index (PRI)</td>
</tr>
<tr>
<td></td>
<td>o Algotect Descriptor Index® (ADI)</td>
</tr>
<tr>
<td></td>
<td>o Mean Word Score (MWS)</td>
</tr>
<tr>
<td>Psychological measures</td>
<td>• Fear of Pain Questionnaire (short form) (FPQ-SF)</td>
</tr>
<tr>
<td></td>
<td>• Anxiety Sensitivity Index (ASI)</td>
</tr>
<tr>
<td></td>
<td>• Pain Anxiety Symptoms Scale (PASS-20)</td>
</tr>
</tbody>
</table>

1. **Cold Pain Threshold (CPT)** at the volar forearm was measured using a peltier thermode (Medoc TSA II, Israel), a device commonly used and considered to be reliable and valid for the assessment of thermal QST (Heldestad et al., 2010; Moloney et al., 2012). The Medoc device is able to reach a minimum temperature of 0°C in contrast to the Somedic thermode that is only able to achieve a minimum of 5°C. The Medoc thermode has been widely used and is accepted as showing good reliability (Moloney et al., 2012).
The 3x3cm contact probe was attached to the forearm test site with a velcro strap. The thermode temperature was set to drop at a constant rate of 1°C/sec from a baseline temperature of 32°C. Standard Method of Limits (Rolke, 2006; Heldestad et al., 2010) and standardised instructions were used. For cold detection threshold subjects were instructed to depress the hand-held switch as soon as they perceived any cooling change from baseline. For cold pain threshold, subjects were instructed to press the switch as soon as the cooling sensation changed to one of painful cold. This temperature was recorded (°C) and the thermode returned to baseline. For both detection and pain threshold tests, one practice was followed by 3 trials, each separated by a randomly assigned pause of between 3 and 6 seconds. The mean of the 3 trials for CPT was calculated for analysis.

2. Sustained cold temperature stimuli were tested at the same test site using the same Medoc TSA II thermode. As for threshold testing, subjects were given several minutes to adapt to the baseline temperature of 32°C. The software then delivered a sustained temperature of either 10°C, 15°C or 20°C, in pre-randomised order over five minutes. This was found to be the maximum possible time period that could be maintained. The 5 minutes included 30 seconds to descend to the target temperature and 30 seconds to return to baseline of 32°C, leaving a total of 4 minutes sustained at the target temperature. A five-minute pause then followed. The next sustained cold temperature was then delivered, followed again by a five-minute rest at baseline temperature. This was repeated a final time for the third sustained temperature.

Response to each sustained cold temperature was measured at 30-second intervals throughout the application. Every 30-seconds, subjects were prompted to rate the intensity of the sensation they were experiencing using the e-VAS. Every 1-minute subjects were also asked to select words to describe the sensation:

a. Intensity of cold, heat, unpleasantness and pain sensations was measured using the newly developed electronic VAS device (Appendix 2). A LabView programme converted output voltage from a series of four calibrated linear potentiometers into values equating to a 100mm visual analogue scale. The programme also provided a timing schedule, with subjects prompted every 30 seconds to provide the next rating by a light on the potentiometer box. Subjects were blind to the actual value of each rating, with the ends of each slider labeled "Minimum imaginable" and "Maximum imaginable" one for each sensation. For each
temperature and each VAS sensation, area under the 5-minute time-VAS curve was calculated, together with maximum VAS.

Figure 5.1: Electronic VAS potentiometer box with sliders for each sensation of cold, heat, unpleasantness and pain

b. Quality of sensation: At 2-minute intervals through the 5-minute application of each sustained temperature, subjects were asked to rate the quality of the sensation they were experiencing. As for Studies 1 and 2, subjects were asked to select words (no minimum or maximum) from the McGill Pain Questionnaire descriptor list, which the investigator recorded. If a subject reported that the sensation was normal this was also recorded but was given no score. The standard MPQ Pain Rating Index and ADI were calculated, as described below, with a higher score in both denoting a more severe response.

3. Skin Temperature Measurement. A digital thermistor probe was placed underneath the edge of the thermode probe, avoiding contact with the thermal plate. The probe was connected to a battery-operated visual display, which provided a reading of temperature in °C. An initial baseline reading of skin temperature was taken before the start of the first sustained cold stimulus, then one minute after each sustained stimulus had finished and the temperature returned to baseline, i.e. at 6 minutes, 16 minutes and 26 minutes.

4. Psychological Measures. Three questionnaires that measure different aspects of fear and anxiety related to pain were also included:

- Fear of Pain Questionnaire (short form) (FPQ-SF): measures fear about situations that would be expected to cause pain. FPQ has been shown to predict experimental pain intensity in normal subjects. A shortened form (20 items) has recently been found to have better internal consistency and construct validity than the full version (Asmundson et al., 2008). The 20 items are scored using a Likert scale (min score 20, maximum score 100).

\(^2\) All psychological questionnaires open access
• **Anxiety Sensitivity Index (ASI):** assesses beliefs that anxiety experiences have harmful personal effects, considered by some to be a more useful tool to measure the effects of anxious behaviour on pain than assessing just frequency of anxiety (Peterson & Heilbroner, 1987). It is a short survey of 16 items, which applies a 5-point Likert scale ranging from 0 (agree very little) to 4 (agree very much), score range from 0 to 64. This scale has been shown in a number of studies to correlate more closely with fear of pain than the conventionally-used Spielberger State-Trait Anxiety Index.

• **Pain Anxiety Symptoms Scale (PASS-20):** is a 20-item questionnaire that assesses pain-related anxiety, using a 6-point Likert scale anchored from 0 (never) to 5 (always) (total score range from 0 to 100). It has been shown to be a robust measure in both clinical and non-clinical populations (Abrams et al., 2007).

**Data Management and Analysis**

Data from intensity and quality rating were managed as follows:

**Intensity:** e-VAS values (0-100) for intensity of cold, heat, unpleasantness and pain every 30-seconds for each 5-minute application were converted from text to excel files and then quantified as:

• Area under the 5-minute VAS-time curve (AUC), as a measure of total intensity;
• Maximum VAS value, as a measure of the greatest intensity experienced;
• Time to onset - the time at which first VAS value >0/100 was recorded;
• Time to peak VAS - the time at which the maximum VAS value was recorded.

**Quality:** MPQ descriptor list choices for quality of sensation experienced every minute for 5 minutes were quantified using two indices:

• Pain Rating Index (PRI): the sum of MPQ ranking values for each different word selected, no single word included more than once (Melzack, 1975).
• Algometry Descriptor Index (ADI): combines a descriptor sub-score with weighted values for VAS intensities. Total ADI maximum score is 9, minimum 0. The descriptor sub-score (MWS) is calculated as the mean of weighted values for each word selected and has a total maximum score of 5, minimum 0.

**CPT data:** A mean CPT temperature value was calculated from the three trials of forearm CPT. In addition to mean CPT value, subjects were divided post-hoc into those with high and low CPT according to CPT < or > 15°C, as rationalised in Chapter 2: 2.5.
Normality testing was carried out using a combination of Shapiro-Wilk tests and visual analysis of distribution. All ADI and PRI data were normally distributed (Shapiro-Wilk p = .066 to .200) meaning that parametric statistics could be applied. The VAS element of ADI, for heat, unpleasantness, and pain required non-parametric analysis. When considered as a single dataset, CPT was not normally distributed and required non-parametric analysis. However, when divided dichotomously according to high or low CPT, data was normally distributed and so parametrics could be used. Data were then analysed using the SPSS statistical package, version 19. Alpha was set at p < .05 and the following analyses were carried out:

<table>
<thead>
<tr>
<th>Hypotheses and Related Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The ADI will discriminate between three sustained cold temperatures (10°C, 15°C, 20°C).</td>
<td>Repeated measures ANOVA Or Friedman’s Two-Way ANOVA</td>
</tr>
<tr>
<td>• Area under VAS-time curve for cold, heat, unpleasantness and pain</td>
<td>Descriptive analysis</td>
</tr>
<tr>
<td>• Maximum VAS rating: cold, heat, unpleasantness, pain</td>
<td>Pearson’s or Spearman’s Correlation Coefficient</td>
</tr>
<tr>
<td>• Time course: cold, heat, unpleasantness, pain</td>
<td></td>
</tr>
<tr>
<td>• ADI and MWS sub-score descriptor ratings</td>
<td></td>
</tr>
<tr>
<td>• Correlations between temperatures for ADI, MWS and VAS</td>
<td></td>
</tr>
<tr>
<td>2. ADI will be more accurate than the PRI in discriminating between temperatures:</td>
<td>ROC curves to predict CPT group</td>
</tr>
<tr>
<td>• ADI at 10°C, 15°C and 20°C</td>
<td></td>
</tr>
<tr>
<td>• PRI at 10°C, 15°C and 20°C</td>
<td></td>
</tr>
<tr>
<td>3. Associations between cold pain threshold and ADI scores for sustained 10°C, 15°C, 20°C:</td>
<td>Pearson’s or Spearman’s Correlation Coefficient</td>
</tr>
<tr>
<td>• ADI &amp; MWS at 10°C, 15°C and 20°C</td>
<td></td>
</tr>
<tr>
<td>• CPT mean (°C)</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>• CPT grouping (&lt;=15°C)</td>
<td></td>
</tr>
<tr>
<td>4. Similarities in intensity and quality of response to 10°C sustained cold and Gel B (from Study 2).</td>
<td>Descriptive analysis</td>
</tr>
<tr>
<td>5. No association between psychological measures and CPT or ADI at any sustained temperature.</td>
<td>Pearson’s or Spearman’s Correlation Coefficient</td>
</tr>
</tbody>
</table>
5.5 Results

Cohort Characteristics
Twenty-nine pain-free, healthy subjects participated in this study, 22 female and 7 male. Mean age was 38 years (range 19 - 63 years).

Distribution of ADI Data
Figure 5.2 illustrates the normal distribution of the ADI data for two of the sustained temperatures assessed: Shapiro-Wilk analysis for sustained 10°C p = .415; for sustained 15°C p = .474.

![Histograms for ADI data at 10°C and 15°C](image)

**Figure 5.2:** Histograms showing normal distribution curves for ADI data for sustained cold at 10°C and at 15°C.

Hypothesis 1
The ADI scoring system will discriminate between three sustained cold temperatures.

- Total ADI and MWS sub-scores scores were significantly higher during lower temperatures

There was a significant temperature-dependent difference in ADI values, with lower temperature evoking higher ADI total scores: $F_{(2,50)} = 32.44$, $p<.001$. Post-hoc contrasts showed that his difference was significant between each temperature ($p<.001$ for all). There was also a significant temperature-dependent effect for the descriptor-only MWS sub-score: $F_{(2,50)} = 15.04$, $p<.001$. Post-hoc contrasts showed that this difference was significant between 10°C and 15°C and between 10°C and 20°C but not between 15°C and 20°C (Figure 5.3).
• **Individual VAS intensities for cold, heat, unpleasantness and pain were also significantly higher at lower temperatures (Figure 5.4).**

There was also a significant temperature-dependent difference in all VAS sensations for total intensity (AUC): cold $F_{(2,56)} = 34.74$, $p < .001$; heat $\chi^2(2) = 1.12$, $p = .36$; unpleasantness $\chi^2(2) = 35.29$, $p < .001$; pain $\chi^2(2) = 21.59$, $p < .001$. The same pattern was found for maximum VAS: cold $F_{(2,56)} = 32.45$, $p < .001$; heat $\chi^2(2) = 8.86$, $p = .012$; unpleasantness $\chi^2(2) = 24.10$, $p < .001$; pain $\chi^2(2) = 20.63$, $p < .001$

• **There were clear differences in word choice at each sustained temperature (Figure 5.5).**

There was also a temperature-dependent difference in word choice. Cold and icy/freeze were selected by considerably more subjects during the coldest temperature (10°C) whereas 20°C was described as predominantly cool. The coldest temperature of 10°C evoked the highest report of burning as well as the unpleasant dysesthetic words prickling, stinging and tingling. 10°C also evoked the highest report of aching sensation.
Chapter 5  Validation of ADI with thermal stimuli

Figure 5.5: Percentage of subjects choosing the most commonly selected words for each sustained temperature

- **There was no difference between temperatures in time to onset for any VAS sensation and little difference in overall response pattern (Figure 5.6a-d).**

There was no temperature-dependent effect for time to onset with all sensations being experienced immediately at all temperatures: cold $F_{(2,56)} = .221, p = .809$; heat $\chi^2(2) = 1.17, p = .231$; unpleasantness $\chi^2(2) = 7.39, p = .071$; pain $\chi^2(2) = 3.47, p = .342$. Response patterns were also very consistent between temperatures for cold and unpleasantness (Figure 5.6a and c), with peak reached immediately followed by a steady decline in intensity to five minutes. Pain VAS pattern was similar, particularly for low temperature cold (10°C). Heat VAS showed a more erratic pattern for 10°C, vacillating around peak intensity for the first 3.5 minutes, then quickly declining at 3.5 to 4 minutes, before stimulus removal, following by a small increase in intensity again immediately following return of cold stimulus to baseline.

Figure 5.6a-d: Release patterns at each temperature for a) cold, b) heat, c) unpleasantness, d) pain
• There was good internal consistency in ADI response for correlations between sustained temperatures.

There were moderate to strong correlations between temperatures for ADI score between each temperature: 10°C and 15°C: r = .797, p = .001; 15°C and 20°C r = .625, p = .001; 10°C and 20°C r = .489, p = .003.

MWS sub-score showed strong correlation between 10°C and 15°C (r = .819, p < .001), but low correlation between 15°C and 20°C (r = .420, p = .023) and no correlation between 10°C and 20°C (r = .322, p = .088). VAS intensity scores however showed good to strong correlations between all temperatures for each sensation (Table 5.2).

<table>
<thead>
<tr>
<th>VAS</th>
<th>r=</th>
<th>p=</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10°C - 15°C</td>
<td>.831</td>
<td></td>
</tr>
<tr>
<td>15°C - 20°C</td>
<td>.767</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>10°C - 20°C</td>
<td>.684</td>
<td></td>
</tr>
<tr>
<td>Heat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10°C - 15°C</td>
<td>.622</td>
<td></td>
</tr>
<tr>
<td>15°C - 20°C</td>
<td>.693</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>10°C - 20°C</td>
<td>.786</td>
<td></td>
</tr>
<tr>
<td>Unpleasantness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10°C - 15°C</td>
<td>.837</td>
<td></td>
</tr>
<tr>
<td>15°C - 20°C</td>
<td>.806</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>10°C - 20°C</td>
<td>.732</td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10°C - 15°C</td>
<td>.903</td>
<td></td>
</tr>
<tr>
<td>15°C - 20°C</td>
<td>.812</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>10°C - 20°C</td>
<td>.826</td>
<td></td>
</tr>
</tbody>
</table>

**Hypothesis 2**

The ADI scoring system will show greater accuracy than the McGill PRI index in discriminating between temperatures

• There were good correlations between the PRI and ADI indices at each sustained cold temperature (Table 5.3).
Table 5.3: Correlations between PRI score and ADI total and MWS sub-score at each sustained temperature.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>ADI Total score</th>
<th>MWS sub-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>10°C</td>
<td>.691</td>
<td>.676</td>
</tr>
<tr>
<td></td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>15°C</td>
<td>.672</td>
<td>.653</td>
</tr>
<tr>
<td></td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>20°C</td>
<td>.783</td>
<td>.549</td>
</tr>
<tr>
<td></td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
</tbody>
</table>

• Receiver operating curve analysis showed that the ADI had higher sensitivity and specificity than PRI for predicting CPT group.

Receiver operating characteristic (ROC) curve analysis showed that ADI score was considerably better in predicting membership of the CPT cold hyperalgesic group: AUC .832, sensitivity .90 and specificity .74 at a score cut-off of ≥4.9. PRI showed less good predictive capacity (Table 5.4).

Table 5.4: Best-fit ROC curve data for sensitivity and specificity of PRI and ADI indices to predict membership of the CPT >15°C group

<table>
<thead>
<tr>
<th></th>
<th>Area under curve</th>
<th>Cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI</td>
<td>.524</td>
<td>≥4.5</td>
<td>.60</td>
<td>.53</td>
</tr>
<tr>
<td>ADI</td>
<td>.832</td>
<td>≥4.9</td>
<td>.90</td>
<td>.74</td>
</tr>
</tbody>
</table>

Hypothesis 3

There will be significant associations between cold pain threshold value and ADI score at each sustained cold temperature:

• There were positive correlations between CPT and ADI at 10°C and 15°C, driven by MWS descriptor sub-score rather than by VAS ratings.

There were good positive correlations between CPT temperature and ADI score at 10°C (r = .555, p = .002) and at 15°C (r = .567, p = .001), but no significant correlation at 20°C (r = .156, p = .419). When individual components of the ADI were analysed, there were only low positive correlations at 10°C and 15°C only, for heat (r = .431, p = .020), unpleasantness (r = .364, p = .050) and pain (r = .369, p = .049). There were no correlations between CPT and cold VAS at any sustained temperature. However, there
were good to strong correlations between CPT and MWS at 10°C (r=.670, p<.001) and 15°C (r=.546, p=.003), as illustrated in Figure 5.7.

![Figure 5.7: Overlay scatter plot showing the correlation between cold pain threshold and MWS descriptor sub-score at 10°C and at 15°C](image)

- **Those with CPT>15°C showed significantly higher ADI scores than those with low CPT.** There were significant differences in ADI scores between CPT groups at the coldest temperatures (10°C $t_{(28)}=-3.46, p=.002$; 15°C $t_{(28)}=-3.58, p=.001$) but not at the least cold temperature of 20°C ($t_{(28)}=-.820, p=.419$). (Table 5.5).

<table>
<thead>
<tr>
<th>CPT&lt;15°C (n=19)</th>
<th>CPT&gt;15°C (n=10)</th>
<th>$t_{(28)}$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>10°C</strong> ADI</td>
<td>4.21 (0.34)</td>
<td>6.09 (0.37)</td>
<td>-3.46</td>
</tr>
<tr>
<td>MWS</td>
<td>2.26 (0.12)</td>
<td>3.19 (0.15)</td>
<td>-4.69</td>
</tr>
<tr>
<td><strong>15°C</strong> ADI</td>
<td>3.08 (0.40)</td>
<td>5.39 (0.46)</td>
<td>-3.58</td>
</tr>
<tr>
<td>MWS</td>
<td>1.77 (0.16)</td>
<td>2.69 (0.25)</td>
<td>-3.30</td>
</tr>
<tr>
<td><strong>20°C</strong> ADI</td>
<td>2.43 (0.37)</td>
<td>2.97 (0.57)</td>
<td>-.820</td>
</tr>
<tr>
<td>MWS</td>
<td>1.70 (0.20)</td>
<td>1.97 (0.23)</td>
<td>-.862</td>
</tr>
</tbody>
</table>

MWS descriptor sub-score of ADI showed a similar pattern of results: significantly higher for those with CPT>15°C at sustained 10°C and 15°C, but not at 20°C (Table 5.5).

The VAS ratings component of the ADI score for each CPT group showed that heat, unpleasantness and pain were rated significantly higher by those with CPT>15°C at 10°C (heat $t_{(28)}=-2.27, p=.031$; unpleasantness $t_{(28)}=-2.67, p=.013$; pain $t_{(28)}=-2.06, p=.049$). At 15°C heat and unpleasantness were also rated significantly higher by the high CPT group (heat $t_{(28)}=-2.56, p=.016$; unpleasantness $t_{(28)}=-3.23, p=.024$). Pain VAS did
not differ significantly between CPT groups at 15°C: \( t_{(20)} = -1.92, p = .066 \). There were no significant group differences at 20°C and cold VAS did not significantly differ between CPT groups at any temperature (Figure 5.8).

![Figure 5.8: Differences in VAS ratings for cold, heat, unpleasant and pain at each sustained cold temperature between CPT groups.](image)

**Hypothesis 4**

*There will be similarities in intensity and quality of response to 10°C sustained cold and menthol Gel B (from Study 2).*

- **Total ADI scores for sustained 10°C and Gel B were almost identical (Table 5.6).**

<table>
<thead>
<tr>
<th>MWS sub-score</th>
<th>ADI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td><strong>Mean (SD)</strong></td>
</tr>
<tr>
<td>10°C</td>
<td>2.6(0.7)</td>
</tr>
<tr>
<td>Gel B</td>
<td>2.7(0.7)</td>
</tr>
</tbody>
</table>

- **Choice of descriptors was also very similar**

The MWS descriptor sub-score of the ADI showed clear similarities between the sensory experience at 10°C and during Gel B application (mean (SD) MWS scores: 10°C 2.6 (0.7); Gel B 2.7 (0.7)). Specific words choice was also very similar (Figure 5.9). 10°C evoked cold, prickling, numb and achings in greater numbers of subjects and Gel B evoked more burning. Otherwise percentages were equivalent.
Figure 5.9: Percentage choice of key descriptors during 10°C sustained cold and menthol Gel B.

- **There were strong similarities between 10°C and Gel B for percentage of subjects reporting any VAS sensation, although VAS intensity for 10°C was higher for each sensation.**

The percentage of subjects reporting any VAS sensation (VAS>0) for cold, unpleasantness and pain were very similar. Heat was experienced by almost double the number of subjects for Gel B than for sustained cold (51.9% versus 24.1%). However, maximum VAS rating was higher during 10°C cold for each sensation (Figure 5.10).

Figure 5.10: Comparison between sustained 10°C cold and Gel B (Study 2) for percentage of subjects reporting any cold, heat, unpleasantness or pain (primary axis) and maximum VAS rating for each sensation (secondary axis).

- **There were clear differences in response patterns between sustained cold and menthol gel.**

For sustained cold temperature, all VAS sensations reached their peak in the first minute of application, with cold, unpleasantness and pain then declining to removal of the stimulus at 5 minutes. Pain remained constant until sharply declining at 3.5 minutes. For menthol gel however, onset was slower, reaching a peak VAS for cold at around 9 minutes and for all other sensations at around 13 minutes. Decline followed sharply after removal of the gel stimulus at 15 minutes.
Hypothesis 5

*There will be no association between psychological measures and either CPT value or ADI at any sustained temperature.*

- *There were only minimal associations between psychological variables and any cold response measure.*

There were no correlations between FPQ, ASI or PASS-20 and CPT or any of the descriptors indices at any of the sustained cold temperatures. For VAS intensities, there were only four positive correlations, all of which were low to moderate and involved VAS cold or unpleasantness at 15°C or 20°C (FPQ and VAS cold at 20°C: r=.382, p=.041; PASS-20 and VAS cold at 15°C: r=.496, p=.006; PASS-20 and VAS cold at 20°C: r=.372, p=.047; PASS-20 and VAS unpleasantness at 15°C: r=.460, p=.012). There were no correlations with the coldest stimulus of 10°C.

When the psychological data of those with CPT< or >15°C were compared, no group differences were found (FPQ t(28)= .655, p=.518; ASI t(28)= 1.25, p=.224; PASS-20 t(28)= -1.47, p=.154).
• **There was no difference in skin temperature change after each sustained cold temperature**

Skin temperature readings were recorded at baseline and one minute following return to baseline after each sustained cold stimulus. Change from baseline was calculated and compared. Skin temperature reduced on average 1.0 (±0.36)°C following 10°C application, 0.8 (±0.18)°C following 15°C and 0.5 (±0.33)°C following 20°C, however this difference was not quite sufficient to reach significance (F(2,50)=3.25, p=.067). When VAS and descriptor scores were analysed to see if there was any association between skin temperature drop and intensity or type of sensation experienced, only a moderate correlation was seen between MWS sub-score and temperature change after 10°C: r=.416, p=.025. There were no other correlations between any VAS or index scores and temperature change.

### Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Accepted/Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The ADI scoring system will discriminate between three sustained cold temperatures.</td>
<td><strong>Accepted</strong></td>
</tr>
<tr>
<td>2. The ADI scoring system will show greater accuracy than the McGill PRI index in discriminating between temperatures</td>
<td><strong>Partially accepted</strong></td>
</tr>
<tr>
<td>3. Significant associations between CPT and ADI score at each sustained cold temperature:</td>
<td></td>
</tr>
<tr>
<td>a. Positive correlations between CPT and ADI</td>
<td><strong>Partially accepted</strong></td>
</tr>
<tr>
<td>b. Those with CPT&gt;15°C will show significantly higher ADI for each sustained temperature</td>
<td><strong>Partially accepted</strong></td>
</tr>
<tr>
<td>4. Similarities in intensity and quality of response to 10°C sustained cold and Gel B (from Study 2).</td>
<td><strong>Accepted</strong></td>
</tr>
<tr>
<td>5. No association between psychological measures and either CPT value or ADI at any sustained temperature.</td>
<td><strong>Accepted</strong></td>
</tr>
</tbody>
</table>
5.6 Discussion

The ADI scoring system was developed as a result of findings from Studies 1 and 2 that a measure of response to cold that combined intensity and quality elements was able to discriminate between graded menthol-cold stimuli. This study therefore sought to validate the newly develop ADI scoring system within the context of conventional thermal-cold stimuli.

Discriminative ability

The study found that the ADI was able to discriminate clearly between graded cold temperatures, showing a significant difference in total score between each of the three sustained temperatures (p<.001 for all). There were also good correlations between temperatures for ADI score (r = .797 to r = .489), suggesting that individuals responded consistently to the stimulus, in a graded manner. This suggests that the ADI exhibits good internal construct validity.

The ADI is composed of intensity and quality elements and these were also separately analysed to explore whether temperature-dependent effects were similar across both domains. VAS intensity showed significant temperature-dependent effects, with VAS cold, heat, unpleasantness and pain intensity all significantly higher for lower sustained temperatures (Figure 5.4). There were also strong correlations between each temperature for each VAS sensation indicating that subjects responded consistently to the different temperature stimuli but with graded levels of intensity. In addition to providing evidence of construct validity, this finding supports the proposal that decreasing cold temperature is signaled in part by increases in ‘volume’ which may reflect increased recruitment of TRPM8 channels and increased calcium influx into cold-sensitive Aδ-fibres to activate increasing numbers of action potentials. TRPM8 is reported to be the only channel that signals cold sensation in both innocuous and noxious contexts (Knowlton 2013) so that increasing cold intensity, even down to 10°C must involve activation of nerve fibres that express TRPM8. In addition to TRPM8 channels, cold intensity is coded by Nav1.8, a sodium channel expressed predominantly on Aδ and c-fibres that co-express nociceptive channels and peptides such as TRPV1 and substance P (Zimmermann et al., 2007). Whereas other sodium channels deactivate in the presence of decreasing temperature, Nav1.8 has been shown to stay open and even increase its firing in response to intense cold. Although Nav1.8 has not been found on cold-sensitive Aδ-thermal fibres, it is found co-expressed with TRPM8 on cold-sensitive Aδ- nociceptors (Knowlton, 2011). This may well be a key mechanism in
signaling cold even at 10°C. The finding that low levels of unpleasantness and pain were also increasingly experienced as temperature reduced, provides further evidence for the activation of lower threshold nociceptors that are cold sensitive. The particularly strong correlations in unpleasantness and pain intensity between 10°C and 15°C suggest that from 15°C and below, greater numbers of cold-sensitive nociceptors are activated. This supports basic science proposals that 15°C is a key temperature in a healthy humans for transition between low threshold, non-noxious signaling Aδ-fibres towards activation of higher threshold nociceptors which more clearly signal a noxious sensation (Campero et al., 2009; Campero & Bostock, 2010).

The descriptor component of the ADI (MWS) was analysed and found to also show significant temperature-dependent effects. In other words, in addition to increases in ‘volume’ of sensations, decreasing temperature caused a change in the quality of sensation experienced. Analysis of individual words selected at each temperature develops this: least cold (20°C) was experienced as a predominantly cool sensation with some warmth associated with the coolness and some cold. In contrast the coldest (10°C), a temperature close to noxious for many healthy subjects, was experienced as icy for 55% of subjects, with clear elements of burning, prickling, stinging and aching sensations. Sensations of cool and cold underlay these additional sensations. However, this change in quality towards more dysesthetic and noxious sensations appears to have started for many subjects at 15°C, suggesting once again that it may be a transition temperature. Total MWS score data supports this: although score difference was significant between both colder temperatures and between 10°C and 20°C (both p<.001), the difference between 15°C and 20°C was not significant (p=.088). There was also strong consistency in response between the two coldest temperatures (r=.819), but poor or no correlation in MWS score between these temperatures and 20°C.

The sensations of icy/freezing, prickling, stinging and burning have been associated in earlier studies with disinhibition of c-fibres following cold stimulation during A-fibre block (Fruhstorfer, 1984; Susser et al., 1998) and in more recent studies with activation of TRPA1 channels (Namer et al., 2005). The role of TRPA1 in signaling noxious cold, particularly in the absence of inflammatory pathology, is disputed (Bautista et al., 2012). However, these channels have been found on low and high threshold cold-sensitive c-fibres (and possibly Aδ nociceptors) where they are co-expressed with Nav1.8 and TRPV1 as well as TRPM8 (Kim et al., 2012). Cinnamaldehyde and allicin which are exclusively TRPA1 agonists evoke sensations of stinging and burning,
although not of cold (Bandell et al., 2004; Namer et al., 2005). Increasing selection of noxious quality sensations during the transition temperature range of 15° to 10°C implies activation of additional nociceptive pathways in addition to cold pathways. The fact that there is no lag time for activation of unpleasantness or pain VAS (Figure 5.6c and d) also suggests that faster conducting Aδ nociceptors are playing a key role, particularly by the time 10°C has been reached.

It has been suggested that paradoxical heat is a clear sign of abnormal sensory interpretation (Maier et al., 2010) probably caused by dorsal horn disinhibition of normal cold inhibitory influence over heat transmission (Susser et al., 1998). However, in this study, heat appears to have been a rather ambivalent, low intensity underlying sensation. The word warmth was reported by 55% of subjects at 20°C but only 21% at 10°C, suggesting that it was more noticeable at less cold temperatures. However heat VAS data found that only 10% experienced any heat sensation at 20°C whilst 20% reported heat VAS>0 at both colder temperatures. Response pattern for heat (Figure 5.6b) was clearly different to that for cold, unpleasantness or pain, with no immediate decline in intensity, although mean intensity level remained very low even at peak (<20/100 VAS) and occurred at the same time as more intense cold. Both heat and cold channels are present on many low threshold Aδ-fibres, meaning that they can signal either sensation. Akopian (2011) has reported that warmth and cold are reciprocally balanced by the influence of protein kinase C and this may explain the apparent co-sensation of both warmth and cold. It may be that a burning sensation is more indicative of changes in sensory processing. At normally noxious temperatures this is likely to be a sign of activation of TRV1 and/or TRPA1 on c-fibre nociceptors. A burning sensation at normally innocuous temperatures may well indicate sensitisation of c-fibres and TRPA1 receptors or dysfunctional spinal disinhibition.

**Criterion validity: CPT**

In addition to being able to demonstrate discriminative ability between graded temperatures, the ADI also showed good associations with CPT, the conventional measure of cold pain sensitivity. ADI correlated moderately well with CPT at 15°C although not at 20°C. The descriptor element of ADI showed stronger correlation with CPT than any VAS values (Figure 5.7), suggesting that cold pain threshold is associated less with increased intensity of pain or unpleasantness and more with a change in the quality of that noxious sensation. This also corresponds with findings from Study 1 where CPT correlated with descriptor score (calculated as McGill PRI) but not with VAS
intensity (Chapter 3, Hypothesis 4). It has been proposed that a CPT > 15°C is a marker for cold hyperalgesia. In the current study therefore, associations between CPT and ADI were also explored by categorising subjects according to CPT < or > 15°C. ADI scores were significantly higher in those categorised as thermal-cold hyperalgesic at the lowest temperatures but not at 20°C (Table 5.5). The pattern was mirrored by both VAS and MWS sub-scores, with VAS heat, unpleasantness and pain increased and MWS scores significantly elevated in those with CPT > 15°C. It is notable that VAS cold intensity was not significantly elevated in those with cold hyperalgesia, indeed for 10°C and 20°C VAS values were very similar (Figure 5.8). This supports the suggestion that noxious cold is more a matter of change in quality than just increased load of the same sensation. Equally it underlines that TRPM8 channels are not deactivated as sensation becomes more noxious, but appear to continue to fire at much the same intensity alongside additional noxious signals. Overall, these findings not only reinforce the likely significance of 15°C as a transitional point for cold response, but also that the ADI scoring system is effective in differentiating between those with and without cold hyperalgesia.

**Criterion validity: PRI**

The discriminatory ability of ADI score was compared with that of the McGill Pain Rating Index (PRI) as an alternative method of evaluating criterion validity. The PRI showed significant similar temperature-dependent effects when measuring response at each sustained temperature, demonstrating that it is a useful tool for measuring changes in sensory or affective quality in experimental pain, as well as changes in spontaneous pain quality for which it was originally designed (Klepac et al., 1981). However, when ROC curve analysis was applied to compare accuracy of the ADI and PRI in predicting CPT group, the sensitivity and specificity values for the ADI were considerably higher than for PRI (Table 5.4). In many ways it is surprising that the PRI predictive capacity for CPT was as high as found. The score does not include any intensity data and is based on words weightings developed from the sensory, affective and evaluative characteristics of spontaneous pain of people with chronic pain disorders (Melzack, 1975). The good overall performance of the PRI in the characterisation of experimental sustained cold stimuli (both in the current study and also in Studies 1 and 2) adds weight to the hypothesis that severe chronic pain experience is strongly associated with sensory processing changes. However, the superior performance of ADI score in predicting CPT group reinforces the value of
combining intensity and weighted descriptor values in a single measure of cold response.

*Comparisons between sustained thermal-cold and sustained menthol-cold*

In order to compare responses between thermal-cold and menthol-cold stimuli, the ADI scoring system was applied post-hoc to Gel B data from Study 2. Non-statistical comparisons between response in the current study to the lowest temperature (10°C) and to the highest menthol concentration from Study 2 (Gel B) showed very clear similarities and several marked differences. ADI scores were remarkably similar both in terms of means and variance (Table 5.6). This was mirrored by strong similarities between MWS descriptor sub-scores. Individual word choices were also strikingly similar. Both cold stimuli evoked predominantly cool and cold sensations, with similar percentages of subjects also selecting icy/freezing. This strongly suggests that the basic mechanism of TRPM8 activation is similar for thermal cold and menthol stimuli. Intensity for maximum cold VAS was slightly higher for sustained 10°C than for the gel. This may be a reflection of the more direct activation mechanism for cold temperature on TRPM8 via a single binding site rather than the four sites for menthol (Janssens & Voets, 2011) or it may reflect the ability of cold temperature to sensitise channels such as Nav1.8 and so increase the initial influx of sodium ions causing a more robust action potential.

The percentage of subjects reporting some VAS unpleasantness and the maximum VAS rating for unpleasantness was also very similar between thermal and menthol cold stimuli and there was close association in choice of mildly unpleasant words such as tingling, stinging and penetrating. The word burning was reported by slightly more subjects during menthol application, but associated with low intensities of pain and unpleasantness. This appears to indicate that there is a similar activation of low threshold nociceptive processes for both cold and menthol stimuli that are largely still at an innocuous level of intensity. As mentioned above, it has been proposed that as cold temperature descends towards a more noxious temperature (perhaps 15°C and below), this is signaled by a change in balance between Aδ thermo-receptor fibres, Aδ nociceptors and then C2 cold and warm sensitive nociceptors. This produces a change in quality of sensation from thermal cold to thermal plus mildly noxious dyasaesthetic sensations such as stinging burning or prickling as nociceptors are activated. The similarity in report of these sensations despite the physiological differences between
cold temperature and menthol reinforces that similar mechanisms beyond just TRPM8 activation are triggered by both stimuli.

Although the percentage of subjects reporting some pain was almost identical, and the level of intensity was still less than 30/100, sustained 10°C cold evoked almost 10% greater pain than menthol Gel B (Figure 5.10). The words pricking and aching were also selected by more subjects for sustained cold than for menthol. Higher intensities for both of these qualities have been reported in previous studies of low temperature noxious cold temperature (Davis & Pope, 2002), suggesting perhaps that even a borderline noxious temperature such as 10°C may activate slightly different pain pathways to menthol. The word aching is often associated with activation of slow conducting c-fibres, although prickle is probably more associated with faster Aδ-nociceptors. The alternative explanation may relate to changes in skin temperature during sustained cooling. It has been previously shown that sustained menthol does not alter skin temperature, despite cold sensory effects (Binder et al., 2011). Few other studies have investigated sustained cold, due partly to technological constraints. The current study however used a thermistor to measure skin temperature at the start and mid-way between each cold stimulus. There was a clear cooling effect that was greatest following the 10°C stimulus, although not statistically different to other temperature stimuli. Closer review of use of aching at each temperature shows that it tended to be accompanied by the word numb and that both words were used at all temperatures, although increasing in incidence as temperature decreased. Vaso-constriction during skin cooling has been reported to activate vascular nociceptors, which may evoke this aching sensation (Kreh et al., 1984). Laser Doppler studies have reported that vaso-constriction can occur during as little as a 30-second sustained 20°C cold stimulus (Lindblad et al., 1990). It seems likely therefore that the aching, numbness and possibly pricking felt during 10°C cold was a reflection of thermal change effects in vascular nociceptors. This extra component of c-fibre input may also explain the higher intensity pain experienced.

Despite such similar overall responses and apparent similarities in mechanisms, the marked difference in response pattern between temperature-cold and menthol-cold seems surprising (Figures 5.11a-b). It suggests that initial transduction mechanisms and deactivation mechanisms are different between the two stimuli, although the overall effect is very similar. Each temperature-cold stimulus peaked immediately for cold, unpleasantness and pain intensity and then steadily declined. In contrast,
menthol-cold showed a lag of about two minutes before gradually increasing to peak at around 9 to 12 minutes and only just starting to decline before removal of menthol at 15 minutes. Thermal cold is able to pass through the skin and epidermis at a much faster rate than a chemical, reaching the TRPM8 channels almost immediately. There is a single activation site for thermal-cold on TRPM8 (Janssens & Voets, 2011), which works in concert with other calcium channels. Lowering of temperature immediately lowers membrane voltage so that calcium is immediately released across the membrane (Babes et al., 2010). Menthol penetrates the epidermis quickly by disrupting the lipid layer of the stratum corneum however this will inevitably require a longer period of time (Obata et al., 2006). TRPM8 is also activated by menthol indirectly, meaning that an additional lag time is likely. Once a menthol molecule binds to its site on the channel a second-messenger cascade is activated, which in turn activates influx of calcium ions and eventual cell activation. TRPM8 has four menthol binding sites that have been shown to gradually fill with menthol molecules, enabling a progressive increase in cell activation, which is likely to be experienced as the gradual increase in intensity of cold to a peak at 9 minutes. Interestingly, this mechanism also suggests that TRPM8-expressing cells may also be predominantly responsible for the unpleasantness and pain sensations since these also showed a very similar initial lag followed by gradual increase to peak at 9 minutes. The steady decline from peak for cold temperature stimuli corresponds with basic data reporting the role of PIP₂ in deactivating TRPM8. As calcium ions increase, PIP₂ is dislodged from its binding site on TRPM8, triggering desensitisation of the channel (McCoy et al., 2011). It is unclear whether the same mechanism controls the eventual decline in intensity from a menthol stimulus, although if so, it would appear to require binding of all four menthol sites before this occurs.

The ADI scoring system has therefore demonstrated the ability to discriminate between graded thermal stimuli both in terms of intensity and quality, demonstrating good internal construct validity. In addition, the ADI scoring system also showed good criterion validity, with good associations both with the PRI score and also with the conventional measure of cold response. The ADI was able to differentiate between normal and more hyperalgesic responses to cold as measured with CPT. These findings suggest that the ADI is an appropriate measure for evaluating cold hyperalgesia.
5.7 Summary
In summary, this study has provided evidence to support the validity of the new menthol test, quantified with the ADI score, as an alternative test for cold hyperalgesia. The ADI score showed very good ability to discriminate between cold temperatures, with responses for both quality and intensity elements consistent at different temperatures. Proposed basic science mechanisms for response to cold were supported by the increases in intensity and change in quality data from this study.

The ADI also showed a good ability to differentiate between those with a more hyperalgesic response, showing that higher CPT was associated with increased intensity of cold, heat, unpleasantness and pain as well as changes to the type of sensation experienced. Once again, this data provides useful psychophysical support for basic science mechanisms.

Good associations between ADI and the similar McGill PRI score during sustained temperatures were shown, with the ADI in fact showing a better predictive ability for CPT group than PRI. The similarities in ADI-measured response between sustained cold and menthol gel also supports the use of menthol, at an appropriate concentration, as an alternative sustained cold stimulus.

The menthol test next needs to be assessed for test-retest reliability, before being applied alongside CPT and other measures of nociception and pain in participants with a chronic pain disorder in order to evaluate its ability to identify those who have clearly discernible signs of hyperalgesia. This will provide final concurrent criterion validity.
Chapter 6

Study Four

Evaluation of test-retest reliability

4.1 Abstract

Background and Aims

Reliability is a key component of any diagnostic test. Previous studies have suggested that thermode-cold testing may not always be consistent. Although the menthol test has demonstrated good validity and internal reliability, its repeatability over several test sessions is yet to be evaluated. This study therefore sought to evaluate the intra-rater test-retest reliability of the menthol test and sub-components of the ADI scoring system for healthy subjects over two test occasions.

Method

Twenty-six healthy volunteers (nine male and 17 female) experienced a menthol test application for 15 minutes on two test occasions separated by at least 24 hours. Test environment and rater was the same on each occasion. The same volar forearm site and application method was used as in previous studies (Chapters 3 and 4). The same response measurement method was also used: VAS intensity ratings for cold, heat, unpleasantness and pain were taken every minute and descriptors chosen from a McGill Pain Questionnaire descriptors list every two minutes. For each test session Algotect Descriptor Index (ADI) total score was calculated. Intra-Class Correlation Coefficients (ICC) were calculated for response reliability between sessions for ADI, each VAS intensity variable and for Mean Word Selected (MWS) descriptor sub-score. Gender differences were also evaluated.

Results

Total ADI showed excellent levels of test-retest reliability with an ICC value of .945 (95% CI: .878 - .975). Sensation quality and intensity sub-components also showed high reliability. MWS descriptor sub-score showed similarly excellent reliability: ICC = .938 (95% CI: .862 - .972). Individual word choice was very closely matched from one session to the next for both noxious words such as burning or stinging and innocuous words such as cool or tingling. There was a difference between sessions for the thermal hot words warm and heat. VAS ratings showed a tendency towards reduction in intensity at the second test session. Although excellent reliability was still demonstrated (ICC = .931 - .886), 95% CIs were larger for heat and unpleasantness,
indicating greater variability between individuals. There was no significant gender
difference in total ADI score. A non-significant trend towards higher pain VAS for
females and higher heat and unpleasantness VAS for males were the only intensity or
quality sub-component gender differences.

**Conclusion**
The menthol test scored with the ADI evokes a highly reliable response in healthy
adults across two test sessions. Future studies evaluating intra-rater reliability in larger
cohorts would be valuable.
6.2 Introduction & Background

Chapters 3, 4 and 5 demonstrated that, at a defined concentration and using the ADI scoring system, the topical menthol test exhibited good levels of validity as an assessment tool for response to cold. Chapters 3 and 4 showed the dose-dependent sensory effects evoked by graded concentrations, both in terms of intensity and quality of sensation experienced. Chapter 3 also showed that the menthol test correlated well with cold pain threshold (CPT) values, the conventional method for assessing cold response. Although this initial study revealed the potential problems with application consistency with a liquid formulation (Chapter 3), subsequent testing of an equivalent gel formulation demonstrated good response consistency and a slightly improved response time (Chapter 4). The gel study indicated that combining the components of sensation intensity and descriptive quality provided better discriminative sensitivity.

The higher concentration of menthol demonstrated the greatest sensitivity in identifying those with an atypical cold response. Chapter 5 sought to validate the new response scoring system developed from findings in Chapters 3 and 4. The Algotect Descriptor Index (ADI) scoring system was applied to three thermal-cold stimuli and showed good temperature-dependent differences in score. This dose-dependent effect was demonstrated for both intensity and quality sub-components of the score, thereby illustrating the good internal consistency of the measure. Criterion validity was demonstrated by comparison to conventional CPT values, with the ADI showing high levels of sensitivity and specificity in predicting CPT-cold hyperalgesic group membership.

In addition to validity it is important that a test also demonstrates good reliability, since a test which of itself varies wildly from one occasion to another, cannot be considered a useful measurement tool. A reliable measure is one that shows acceptable levels of measurement error and that shows consistency between test occasions and between raters. In the case of the menthol test, it is important to evaluate whether an individual who scores as cold hyperalgesic or non-hyperalgesic on one occasion, scores similarly on a second test occasion. In addition, it is valuable to note the consistency of sensation experienced, particularly in terms of quality, as the current study has proposed that the quality of sensation reported is important to determine a hyperalgesic response.

There are a number of inter-linked factors involved in the assessment of reliability for the menthol test. There are factors relating to the reliability of the gel stimulus
provided to the individual. This may be influenced by ingredient and potential preparation method variability, by storage conditions and by specific application method. Many of these factors have already been evaluated in small pilot studies (Appendices 1 and 2) and Chapters 3 and 4 demonstrated good internal consistency of response between concentrations, suggesting that the formulation itself was relatively stable.

An important factor relates to the stability of the phenomenon of cold response itself. If a normal individual’s response to the same cold stimulus varies from day to day then no matter how reliable the menthol test is per se, a result on a single occasion may not be a reliable reflection of whether they are cold hyperalgesic or not. As discussed in Chapter 2: 2.5, there are a number of methods for assessing cold response, each of which have been shown to have questionable reliability. The cold pressor test is widely used as an inexpensive method for assessing a cold hyperalgesic response. However, it has been reported that the test has questionable test-retest reliability unless a sophisticated temperature control system is applied to ensure a completely consistent cold stimulus (Mitchell et al., 2004). In addition, cold pressor result has been associated with psychological factors, suggesting that its reliability may be influenced by variations in an individual’s mood more than with other QST tests.

Cold pain threshold is more often used in research contexts to assess cold pain response. Although there can be problems with equipment, methodology and statistical approach, a recent systematic review concluded that CPT overall demonstrates fair to good reliability (Moloney et al., 2012). A number of studies have reported an inability to analyse CPT measurements due to lack of data. A recent study of QST test-retest reliability in OA patients by Wylde et al. (2011) was unable to analyse results for CPT due to insufficient data. The 5°C cut-off for the MSA (Somedic) thermode resulted in too many subjects having no CPT value recorded. A number of studies have successfully reported test-retest reliability for CPT, although lack of consistency in method and statistical analysis approach can make it difficult to generalise with confidence. Heldestad et al. (2010) used Coefficient of Repeatability (CR) to report good to excellent reliability for CPT (using Method of Limits), both between days and between sessions within the same day. The majority of other studies have used Intra-Class Correlation Coefficients, making them easier to compare. Sand et al. (2008a) reported a mean ICC of .55 for CPT across all head, neck and thoracic sites when tested twice over three to ten days. Wasner and Brock (2008) reported that CPT re-tested over two days
resulted in an ICC of .948, although this reduced to ICC= .781 when retested over 21
days. High levels of one-week test-retest reliability (ICC=.855) for multiple body sites
in patients with a range of neuropathic pain disorders have also been shown by Geber
et al. (2011). Moloney et al. (2011) assessed test-retest reliability of thermal QST on the
hand and reported that, although CPT showed high ICC (.87 - .94), inter-subject
variability was also high (Coefficient of Variance: 84.9-90.2%), leading the authors to
conclude that CPT is valuable for group analysis but may be less reliable for individual
analysis. However, the high variability in this study involved actual differences of less
than 2°C. Given that healthy inter-subject variability can be up to 15°C and within the
context of identifying a normal versus abnormal response, a 2°C difference is unlikely
to be clinically or diagnostically significant.

There is some debate over whether CPT is as variable between test sites as PPT or HPT
have been shown to be. Theoretically, if CPT is a predominantly centrally-controlled
phenomenon, inter-site variability should be small. Although several studies report
statistically significant differences in CPT temperatures between sites (Sand et al.,
2008a; Stone et al., 2012), the actual difference is often only 1°C or so (Rolke et al.,
2006). As noted above, it is reasonable to assume that such a small difference is not
clinically significant.

The generally good results for thermode-cold testing reliability suggest that cold pain
response is a relatively stable psychophysical phenomenon. Although previous studies
in this thesis have demonstrated a good relationship between CPT temperature and
response to the menthol test, there are fundamental differences in modality and
measurement concept between the two tests. Whereas CPT quantifies cold response as
a single temperature at which an individual experiences cold pain, the ADI quantifies
response to the stimulus in terms of intensity and quality description of the sensation.
Grafton et al. (2005) evaluated the test-retest reliability of the Short Form McGill Pain
Questionnaire at two time-points in 71 subjects with knee or hip osteoarthritis. The
sensory element of the descriptors section of the MPQ showed an ICC of .95 (95% CI:
0.92-0.97) and a standard error of measurement of 1.64. A recently completed study
by Whitnell et al. (Unpublished) is the only one to evaluate reliability of CPT using
quality descriptors. This study assessed CPT using conventional methodology at three
upper limb and lower limb tests sites and then asked subjects to select words from the
McGill list to describe the sensation. They found that CPT reliability between three test
sessions, each separated by at least 24 hours, was high (ICC=.87 - .91), with a low mean
CV of 31.8%. In addition, index scores for descriptor choice at CPT also showed high ICC values, with the Pain Rating Index (PRI) showing an ICC= .850. Despite clear inter-subject variability in word choice, there was high intra-subject consistency in that choice between each session. This study demonstrates that the phenomenon of response to cold is consistent, whether measured as a CPT value or quantified using descriptor indices.

There is no standard method for determining reliability, although ICCs appear to be the most widely used. However, as demonstrated by several studies, reliance on only ICC values can be misleading, since it has been reported that when there are either very large or very small variations in data the ICC may over-estimate or under-estimate reliability (Portney & Watkins, 2008). It has been suggested that for a measure such as the McGill pain questionnaire that alongside reliability values there should be some measure of precision, such as 95% confidence intervals or standard error of measurement (Melzack, 1975; Grafton et al., 2005).

No studies to date have considered the between-sessions reliability of topical menthol as an assessment of cold response. The aim of this study therefore was to assess the test-retest reliability of response to the menthol test as measured with the ADI scoring system. Healthy subjects were assessed over two separate occasions, using the same testing method, same rater and in the same laboratory environment.
6.3 Hypotheses

1. There will be good (intra-rater) test-retest reliability for menthol test ADI score between two test occasions.

2. There will be good test-retest reliability for menthol test MWS descriptor sub-score between two test occasions.

3. There will be good test-retest reliability for VAS intensity values for cold, heat, unpleasantness and pain between two test occasions.

4. There will be good test-retest reliability for time to onset and time to peak sensation between two test occasions.

5. There will be no significant difference in ADI, MWS or VAS values between males and females on either test occasion.
6.4 Methods

Subjects
Twenty-six pain-free healthy adult subjects were voluntarily recruited for this study from the Perth community, using word of mouth. Previous studies in this investigation have shown that a sample size of between 25 and 30 participants provides sufficient statistical power to calculate within-subject differences in VAS and descriptor outcome measures. Inclusion and exclusion criteria were as in Table 6.1.

| Inclusion: | Aged 18-70
|           | Able to read and understand English |
| Exclusion: | Current pain >1/10 VAS
|           | Currently taking any analgesic or anti-inflammatory medication
|           | History of systemic inflammatory conditions;
|           | Neurological deficits (motor, cognitive or sensory);  
|           | History of other chronic pain disorders (e.g. fibromyalgia, chronic low back pain);
|           | Skin allergies;
|           | Allergy to menthol or ethanol. |

All participants provided written informed consent before participating in the study. Ethical approval was provided by Curtin University Human Research Ethics Committee (Approval number PT0188/2012 (B)).

Study design and procedures
The study followed a simple test-retest design, with subjects experiencing the same gel formulation on two test occasions, separated by at least 24 hours. A gap of at least 24 hours was chosen since previous small-scale observational work showed that any residual effects of menthol (visual or sensory) had dissipated two hours post gel removal. A previous study of a range of QST outcomes also showed that reliability was similar between same day and two-day testing-retesting (r = .83** or r = .85**).

All testing occurred in a room with the ambient temperature maintained at 23-24°C. The same room and conditions were used on each test occasion for each subject. All menthol testing was performed at the same 2x3cm mid-line volar forearm site, 5cms from the wrist crease. The dominant arm was used for all subjects. Standard arm
positioning was used, with the subjects maintaining their forearm in full supination resting on a table. Before each test, the site was gently cleaned with tepid water and dried with a paper towel. Before menthol testing, subjects were given 5 minutes to familiarise themselves with the McGill (MPQ) descriptor list.

**Preparation of menthol gel:**
As for all previous menthol studies, all menthol formulation preparation was carried out under Pharmacy Lab conditions, using standardised procedures. This study used the Gel B formulation assessed in Study 2 (Chapter 4). The same batch of menthol gel formulation was used for all test-retest trials on all subjects. Gel formulation was stored in opaque glass lidded containers in stable temperature conditions. The gel was brought to room temperature before applying to the subject each time.

**Outcome Measures**
The same measurement method was used to assess response to the gel as was used in the ADI validation study (Chapter 5), using VAS ratings for intensity of cold, heat, unpleasantness and pain and MPQ descriptors list for sensation quality. Intensity ratings were collected every minute and descriptors every two minutes.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable measured</th>
</tr>
</thead>
</table>
| Response to menthol gel formulations B | • Algotect Descriptor Index (ADI):  
  o Intensity of cold, heat, unpleasantness and pain (VAS /100);  
  o Quality of sensation: Mean Word Score (MWS) sub-score. |

**Menthol test application procedure:** Once the test area on the forearm was marked and gently cleaned with hypoallergenic soap and tepid water, 2ml of menthol gel was applied to the site using a 5ml syringe. A Tegaderm dressing was immediately placed over the gel, and the gel then gently spread through the dressing so that it filled the 2x3cm window. A timer was then started for the 15-minute application.

**Intensity ratings:** At one-minute intervals through the 15-minute application, subjects were asked to rate the intensity of the sensation they were experiencing. Intensity of cold, heat, unpleasantness and pain were rated using 100mm VAS scales, each marked “maximum (sensation)
imaginary" and "minimum (sensation) imaginable". Each VAS rating was measured and the value recorded for each sensation at each minute. A final VAS recording was taken at 20 minutes, five minutes after menthol removal.

Quality ratings At two-minute intervals through the 15-minute application, subjects were asked to rate the quality of the sensation they were experiencing. As for previous studies, subjects were asked to select words (no minimum or maximum) from the MPQ descriptor list, which the investigator recorded. If a subject reported that the sensation was normal this was also recorded but was given no score. The ADI was calculated, as described below, with a higher score denoting a more severe response.

After 15 minutes, the Tegaderm dressing was removed and the menthol solution immediately washed from the skin using hypoallergenic fragrance-free soap and tepid water, gently patted dry with paper towels. A further recording of VAS intensities and descriptors was taken at 20 minutes.

Data Management and Analysis

Data from intensity and quality rating were managed as follows:

Algotect Descriptor Index (ADI) was calculated as the sum of 2 components, with a maximum possible ADI score of 9.

1. **Quality: Mean Word Score (MWS):** The mean of weighted values for each selected descriptor word chosen.
   Maximum possible score: 5

2. **Intensity: Max VAS for cold, heat, unpleasantness and pain:** The sum of a weighted values for maximum VAS value for each sensation during application.
   Maximum possible score: 4.

Reliability Analysis

Although the data that results from calculation of the ADI uses a weighted scoring system, it was regarded as interval level data for the current analysis of reliability. Although this is open to debate, previous studies of the McGill MPQ, a similarly
weighted pain scoring system, have proposed that it is appropriate to view this type of
data as interval because increases in score are interpreted as linear and of equal
incremental value (Grafton, 2005). In the same way, with the ADI a score of 6/9 would
be interpreted as twice as severe as a score of 3/9.

Intra-Class Correlation Coefficients (ICC) were selected for the analysis of test-retest
reliability. Although there is a lack of consistency in the literature about the optimal
reliability analysis, ICCs are the most widely accepted and this allows for future
comparison with other studies (Grafton, 2005; Moloney). However, the disadvantage of
using an ICC only is that it provides no indication of the actual amount of error involved
or whether there is a systematic bias. Bland and Altman (1986) suggest that
presentation of additional descriptive data such as standard error of the measurement
(SEM) and scatter graphs so that the reliability results are more easily interpreted.

Normality testing was carried out using a combination of Shapiro-Wilk tests and visual
analysis of distribution. ADI, MWS and cold VAS data were normally distributed
without transformation (Shapiro-Wilk p = .109 to .474). Unpleasantness and pain VAS
data was not normally distributed and so required non-parametric analysis. Data were
analysed using the SPSS statistical package, version 19 with Alpha set at p < .05.

<table>
<thead>
<tr>
<th>Hypotheses and Related Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test-retest reliability for:</td>
<td></td>
</tr>
<tr>
<td>1. ADI score: total</td>
<td><em>Intra-Class Correlation</em></td>
</tr>
<tr>
<td>2. MWS quality sub-score</td>
<td><em>Coefficients</em></td>
</tr>
<tr>
<td>3. VAS intensity: cold, heat, unpleasantness, pain:</td>
<td><em>(95% Confidence Intervals)</em></td>
</tr>
<tr>
<td>• Area under curve</td>
<td><em>Descriptive analysis</em></td>
</tr>
<tr>
<td>• Maximum VAS value</td>
<td></td>
</tr>
<tr>
<td>4. Temporal characteristics:</td>
<td></td>
</tr>
<tr>
<td>• Time to 1st sensation (mins) for cold, heat, unpleasantness and pain</td>
<td><em>Intra-Class Correlation</em></td>
</tr>
<tr>
<td>• Time to peak sensation (mins) for cold, heat, unpleasantness and pain</td>
<td><em>Coefficients</em></td>
</tr>
<tr>
<td>• Time-course for each VAS sensation and key quality descriptors</td>
<td><em>(95% Confidence Intervals)</em></td>
</tr>
<tr>
<td>5. Differences between males and females:</td>
<td><em>Independent t-tests or</em></td>
</tr>
<tr>
<td>Measure</td>
<td>Method</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>ADI total score</td>
<td>Mann-Whitney U tests</td>
</tr>
<tr>
<td>MWS sub-score</td>
<td></td>
</tr>
<tr>
<td>Area under curve for VAS cold, heat, unpleasantness, pain</td>
<td></td>
</tr>
<tr>
<td>Maximum VAS for cold, heat, unpleasantness, pain</td>
<td></td>
</tr>
</tbody>
</table>

* For each ICC calculation a two-factor, mixed effects model was used.
6.5 Results

Cohort Characteristics

26 healthy subjects participated in this study, 9 males and 17 females. The mean age of the group was 48 years (range 33-72).

Table 6.2: Age characteristics of male and female subjects

<table>
<thead>
<tr>
<th>Age (range) (years)</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-29</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30-39</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>40-49</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>50-59</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>60-69</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>70-79</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9</strong></td>
<td><strong>17</strong></td>
<td><strong>26</strong></td>
</tr>
</tbody>
</table>

Distribution of ADI Data

Shapiro-Wilk analysis showed that menthol ADI data was normally distributed on both test days (Day 1 p=.415; Day 2 p=.474) (Figure 6.1).

Figure 6.1: Histograms showing normal distribution curves for menthol ADI data on Day 1 and Day 2 test sessions.

Hypothesis 1

There will be good (intra-rater) test-retest reliability for menthol test ADI score between two test occasions.

- The ADI total score showed high levels of reliability

Intra-Class Correlation Coefficient analysis showed high levels of reliability for the ADI total score between two test occasions: ICC=.945 with 95% CI (.878-.975) (Table 6.3). The mean difference between test days was .25, with the mean score on Day 2 being
5.8% lower than on Day 1 (Figure 6.1a). Scatterplots illustrate the correlation between ADI score on each test occasion (Figure 6.1b).

Table 6.3: Intra-Class Correlation Coefficient (95% confidence interval) for ADI total score and MWS and VAS sub-scores

<table>
<thead>
<tr>
<th></th>
<th>ICC</th>
<th>95% CI</th>
<th>Mean Difference (Day 1-Day2)</th>
<th>Standard Error of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADI total score</td>
<td>.945</td>
<td>.878 -.975</td>
<td>.25</td>
<td>.129</td>
</tr>
<tr>
<td>MWS score</td>
<td>.938</td>
<td>.862 -.972</td>
<td>.11</td>
<td>.082</td>
</tr>
<tr>
<td>VAS Cold</td>
<td>.931</td>
<td>.847 -.969</td>
<td>2.8</td>
<td>2.26</td>
</tr>
<tr>
<td>Heat</td>
<td>.894</td>
<td>.763 -.952</td>
<td>4.4</td>
<td>3.24</td>
</tr>
<tr>
<td>Unpl</td>
<td>.886</td>
<td>.746 -.949</td>
<td>5.1</td>
<td>3.65</td>
</tr>
<tr>
<td>Pain</td>
<td>.928</td>
<td>.840 -.968</td>
<td>2.4</td>
<td>1.45</td>
</tr>
</tbody>
</table>

**Hypothesis 2**

*There will be good (intra-rater) test-retest reliability for menthol test MWS sub-score between two test occasions.*

- The quality component of the ADI showed equally good reliability between 2 test occasions

ICC analysis showed similarly high levels of reliability for the MWS descriptor sub-score: ICC = .938 with 95% CI (.862-.972) (Table 6.3). The mean difference between test days was .25, with the mean score on Day 2 also lower than that on Day 1 but only by 4.2% (Figure 6.2a). Figure 6.2b illustrates the close correlation in MWS scores between the two test days.

![Figure 6.2a](image)

**Figure 6.2a:** Mean (SEM) values for total ADI score and MWS descriptor sub-score for Day 1 and Day 2 test sessions.
• **Individual word choice showed that warm and hot were selected by considerably fewer subjects at Day 2 but that otherwise there was strong similarity in word choice between test days.**

Figure 6.3 illustrates the percentage of subjects selecting specific words on each test occasion. The heat words warm and hot were used by 30% and 34% fewer subjects on Day 2 and fewer subjects chose penetrating. However, burning was selected by a similar percentage of subjects (Day 1: 19.1%; Day 2: 15.4%). Cold-type words and both mild dyasaesthetic (tingling, itching) and unpleasant dyasaesthetic (prickling and stinging) words were selected by very similar percentages of subjects on both days (Figure 6.4).

---

**Figure 6.2b:** Overlay scatterplot for ADI scores for Day 1 and Day 2 (yellow) and MWS scores for Day 1 and Day 2 (green).

**Figure 6.3:** Percentage of subjects selecting specific words on Day 1 and Day 2
**Hypothesis 3**

There will be good test-retest reliability for VAS intensity values for cold, heat, unpleasantness and pain between two test occasions.

- VAS cold and pain showed equally good reliability to the MWS sub-score. Heat and unpleasantness VAS showed greater change from Day 1 to Day 2, although overall ICC values were still high (Table 6.3).

ICC values for cold and pain VAS showed high reliability: VAS cold ICC = .931 (95% CI: .847 -.969); VAS pain ICC = .928 (95% CI: .840 -.968). 85% of subjects reported some VAS cold intensity (Figure 6.5a) and the mean difference in VAS values between test days was small (2.8/100). Mean difference in values for VAS pain was also very small (2.4/100), with between 34.6% and 30.8% of subjects reporting some pain VAS on each occasion (Figure 6.5b).

**Figure 6.4:** Percentage selection of different types of words on each test occasion

---

**Figure 6.5a:** Percentage of subjects rating VAS >0 on Days 1 and 2 for cold, heat, unpleasantness and pain.

**Figure 6.5b:** Mean (SEM) VAS ratings for those subjects scoring VAS >0 on test Days 1 and 2 for cold, heat, unpleasantness and pain.
ICC values for heat and unpleasantness were slightly less good than for cold and pain, with larger 95% confidence intervals: VAS heat ICC=.894 (95% CI: .763 - .952). VAS unpleasantness ICC=.886 (95% CI: .746 - .949). Figure 6.5a shows the percentage of subjects reporting any heat or unpleasantness on each test occasion. Although for those who rated VAS intensity >0 there was a reduction in heat and unpleasantness value by Day 2 (Figure 6.5b), the mean difference in values for the whole cohort between Days 1 and 2, including those who scored VAS 0 on both occasions, was only 4.4/100 and 5.1/100 for heat and unpleasantness respectively (Table 6.3).

Scatterplots illustrate that the correlation for cold VAS and heat VAS between Days 1 and 2 were relatively good (Figure 6.6a) whereas those for unpleasantness and pain (Figure 6.6b) were less close. For pain VAS this lower correlation was caused by the large number of 0 scores.

**Figure 6.6a:** Overlay scatterplot for cold VAS ratings (blue) for Day 1 and Day 2 and heat VAS ratings (red) for Day 1 and Day 2.

**Figure 6.6b:** Overlay scatterplot for unpleasantness VAS ratings (green) for Day 1 and Day 2 and pain VAS ratings (orange) for Day 1 and Day 2.

**Hypothesis 4**

*There will be good test-retest reliability for time to onset and time to peak sensation between two test occasions.*
• **ICC values were more variable for time to onset, with time to onset for cold VAS showing poor ICC reliability. Time to peak showed high reliability for all sensations (Figure 6.7a-b).**

High ICC values were shown for time to onset of first VAS sensation for heat (ICC= .962, 95% CI: .915 - .983) and good reliability for unpleasantness (ICC= .764, 95% CI: .473 - .894) and pain (pain ICC= .761, 95% CI: .467 - .893). Reliability for TTO for cold was only moderate, with a large confidence interval (ICC= .675, 95% CI: .274 - .854). Time to peak however showed high ICC values for each sensation: cold ICC= .899 (95% CI: .774 - .955); heat ICC= .964 (95% CI: .921 - .984); unpleasantness ICC= .925 (95% CI: .832 - .966); pain ICC= .858 (95% CI: .684 - .936).

![Figure 6.7a-b: Intra-class correlation coefficients Day 1 to 2 for mean times to a) onset and b) peak VAS sensations](image-url)

• **Time-course of sensations during application followed a very similar pattern at Day 1 and Day 2 (Figure 6.8a-b).**

![Graph showing VAS values over time](image-url)
Hypothesis 5
There will be no significant difference in ADI, MWS or VAS values between males and females on either test occasion.

- There were minimal gender differences in ADI, MWS or VAS values on either test day. Independent t-tests for ADI and MWS and Mann-Whitney U tests for VAS data showed no significant differences between males and females for any measure on either day 1 or Day 2 (Figure 6.9 and Table 6.4). There were trends towards females reporting greater pain than males, and males reporting more heat and unpleasantness.

Table 6.4: Means (SD) for male and female ADI, MWS and VAS values on Day 1 and Day 2.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>t(25) or U(25)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Day 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADI total score</td>
<td>4.5 (1.4)</td>
<td>4.2 (1.5)</td>
<td>.413</td>
<td>.683</td>
</tr>
<tr>
<td>MWS sub-score</td>
<td>2.8 (.58)</td>
<td>2.5 (.65)</td>
<td>1.09</td>
<td>.289</td>
</tr>
<tr>
<td>VAS Cold</td>
<td>21.0 (15.2)</td>
<td>29.5 (18.8)</td>
<td>-.973</td>
<td>.331</td>
</tr>
<tr>
<td>Heat</td>
<td>19.218 (2)</td>
<td>11.8 (21.8)</td>
<td>-.133</td>
<td>.185</td>
</tr>
<tr>
<td>Unpleasantness</td>
<td>19.8 (7.6)</td>
<td>17.6 (20.3)</td>
<td>-.421</td>
<td>.674</td>
</tr>
<tr>
<td>Pain</td>
<td>4.8 (6.1)</td>
<td>9.6 (14.7)</td>
<td>-.308</td>
<td>.758</td>
</tr>
<tr>
<td><strong>Day 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADI total score</td>
<td>4.2 (1.5)</td>
<td>4.0 (1.8)</td>
<td>.332</td>
<td>.743</td>
</tr>
<tr>
<td>MWS sub-score</td>
<td>2.6 (.62)</td>
<td>2.4 (.84)</td>
<td>.677</td>
<td>.505</td>
</tr>
<tr>
<td>VAS Cold</td>
<td>24.0 (14.6)</td>
<td>23.6 (18.1)</td>
<td>-.081</td>
<td>.935</td>
</tr>
<tr>
<td>Heat</td>
<td>12.6 (11.7)</td>
<td>8.5 (15.8)</td>
<td>-.120</td>
<td>.229</td>
</tr>
<tr>
<td>Unpleasantness</td>
<td>13.7 (11.4)</td>
<td>13.1 (15.8)</td>
<td>-.472</td>
<td>.637</td>
</tr>
<tr>
<td>Pain</td>
<td>1.1 (3.3)</td>
<td>7.9 (10.9)</td>
<td>-.168</td>
<td>.093</td>
</tr>
</tbody>
</table>
**Figure 6.9**: Mean (SEM) VAS ratings for cold, heat, unpleasantness and pain on Day 1 and Day 2, comparing male and female subjects.

### Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>There will be good (intra-rater) test-retest reliability for menthol test ADI score between two test occasions.</td>
</tr>
<tr>
<td>Accepted</td>
</tr>
<tr>
<td>There will be good test-retest reliability for menthol test MWS descriptor sub-score between two test occasions.</td>
</tr>
<tr>
<td>Accepted</td>
</tr>
<tr>
<td>There will be good test-retest reliability for VAS intensity values for cold, heat, unpleasantness and pain between two test occasions.</td>
</tr>
<tr>
<td>Partially accepted</td>
</tr>
<tr>
<td>There will be good test-retest reliability for time to onset and time to peak sensation between two test occasions.</td>
</tr>
<tr>
<td>Partially accepted</td>
</tr>
<tr>
<td>There will be no significant difference in ADI, MWS or VAS values between males and female on either test occasion.</td>
</tr>
<tr>
<td>Accepted</td>
</tr>
</tbody>
</table>
6.6 Discussion

The menthol test, with the ADI scoring system, has been proposed as a new assessment for response to cold. There is no published data regarding the reliability of response to menthol. This study therefore evaluated the intra-rater test-retest reliability of response to the test by healthy participants over two separate sessions.

*High ADI reliability*

The menthol test carried out at a volar forearm site showed excellent reliability, with an ICC of .945 and narrow 95% confidence interval (Table 6.3). Variability was also low, with standard error of measurement (SEM) of .13 for change between days. Figure 6.1b illustrates that the correlation between ADI score on Day 1 and Day 2 was also extremely high (r= .953). There are no previous reported test-retest studies for topical menthol with which to compare these results. Thermode cold pain testing aims to identify the same phenomenon as the menthol test, and so theoretically should demonstrate similar level of reliability, although the clear differences in modality and measurement approach mean that comparisons must be made with caution.

CPT test-retest reliability has been investigated in several different studies, although results are quite variable. For example, Sand et al. (2008b) investigated CPT reliability over seven days at the thorax, head and neck and found considerable variation between sites, reporting a mean ICC for all sites of .55. In contrast, Wasner and Brock (2008) reported excellent reliability (ICC=.948) for CPT at the hand tested over 2 days, reducing to .781 for test-retest over 21 days. This study also showed that VAS pain ratings taken at CPT were also highly correlated between test sessions. Geber et al. (2011) found that CPT test-retest reliability over two days was as high (ICC .855), using participants with a wide range of neuropathic pain symptoms and a variety of test sites. In a recent systematic review of thermal reliability, Moloney et al. (2011) concluded therefore that overall CPT reliability was fair to good.

Yet high variability in CPT has been previously indicated as a reason for avoiding cold response assessment for central excitability (Stone et al., 2012). In a CPT reliability study of the hand Moloney et al. (2011) agreed that caution must be exercised in application of CPT. Although ICC values in this study were comparable to those reported by Wasner and colleague (2008), CPT values showed high standard deviations and Coefficient of Variability. Moloney et al. (2011) concluded that this intra-subject
variability meant that, although CPT was reliable for group comparisons, it was less suitable for individual analysis.

The current study found considerably lower variability alongside high ICC reliability statistics for the menthol test, suggesting that it may be a more consistent tool to use for identifying cold hyperalgesia. High levels of reliability and low variability may partly be a function of a scoring system that is relatively narrow (minimum 0 to maximum 9), as opposed to CPT, which potentially can vary from 32°C to 0°C in a normal cohort. However, it may also be that the ADI scoring system, which combines two sensory domains to characterise response is more sensitive for the particular phenomenon of cold pain. Heat pain is a far more clear-cut and defined sensation, with low variability both between and within individuals (Moloney et al. 2011). Although ICC values for HPT have been reported to be slightly lower than for CPT, this is likely to be a function of the very low standard deviations. Cold pain in normal individuals appears to be a less defined sensation. Whereas the transduction pathway for heat is very clear, segmented by five temperature-specific TRP channels which mark increasing gradations of temperature, cold pain is less clearly signaled, involving a complex interaction between TRP channels, sodium and potassium channels, intracellular cascades and inhibitory-disinhibitory interactions between nociceptive and non-nociceptive fibres. Cold sensation and pain is therefore less defined than heat pain, meaning that accurate identification of the specific temperature at which the cold becomes painful is likely to be more variable. Theoretically, the menthol test allows for greater accuracy by providing more time to respond and by providing a measurement that combines both intensity and type of sensation. In this way, response to the cold stimulus is characterised with more precision than is possible with a single temperature value. Standard CPT testing can only assume that an individual is actually registering a painful cold temperature because the only data collected is that of temperature value. The menthol scoring system also offers greater internal reliability since intensity ratings that include cold and heat and combined with weighted descriptor which also include cold and heat-type words. This theoretically should translate into improved test-retest reliability and reduced variability for the menthol test.

**High descriptor reliability**

The MWS descriptor sub-component of the ADI also showed similarly high reliability (ICC= .938) and low variability. SEM was very low, although 95% CI for ICC were
slightly greater than for ADI alone (Table 6.3). The McGill Pain Questionnaires descriptor section is the most comparable to the MWS. Test-retest reliability for the descriptor component of the MPQ has been assessed by a small number of studies, although none in the same context as the current study. Grafton et al. (2005) assessed the reliability of the MPQ for spontaneous pain in subjects with knee and hip OA over two separate occasions. This study found that the sensory component of the questionnaire, which is most comparable to the menthol MWS, showed excellent reliability, similar to the current study (ICC= .95), although SEM and Coefficient of Repeatability values in the Grafton study were higher. The authors concluded that this variability is a reflection of the natural changeability in OA pain from one day to the next and does not necessarily reflect poor stability of the MPQ. Whitnell et al. (Unpublished) ¹ assessed test-retest reliability of the MPQ during QST testing in healthy subjects. CPT was assessed over 3 sessions using three upper limb and lower limb test sites. MPQ descriptors were selected to describe the sensation at CPT and PRI index used for analysis. This study found that both CPT and PRI scores showed high reliability over time: CPT ICCs= .87 to .91 and PRI ICC= .85. This matches the results from the current study.

Both the study by Whitnell and colleagues and the current study analysed in more detail specific word choice at each time point. Whitnell et al. reported a similar spread of word choice between individuals, with some selecting mild cooling words and others choosing more intense words such as burning or prickling. Yet the intra-subject correlation in words between test sessions was marked, with very little variability (Figure 6.10).

![Figure 6.10](image)

**Figure 6.10**: Percentage of subjects reporting different descriptor words at CPT, comparing three separate test sessions.

*Reproduced from Whitnell et al. (unpublished), with permission.*

---

¹ Whitnell et al (unpublished): abstract in Appendix 5
Chapter 6

Test-retest reliability

The current study also found consistency in word choice for menthol sensation between the two test occasions (Figure 6.3). As with Whitnell et al., this consistency was found both for innocuous words such as cool, cold or tingling and for more intense words such as burning, icy or stinging. This finding strongly suggests that a cold stimulus activates a distinctive neurosensory response in individuals, which remains stable over time. It may be hypothesised that those with the less intense cool response were likely to be activating predominantly Aδ thermal fibres, whereas those who reported more of a burning or stinging sensation were likely to have activated Aδ or c-nociceptors in response to the same stimulus. The possible mechanisms for this in healthy individuals such as altered descending inhibitory drive have been discussed elsewhere. However, the consistency in noxious or innocuous response found in the current study reinforces the idea of an established, centrally-driven mechanism.

The current study did however find an anomaly in the otherwise reliable choice of words across time, with warm and hot words selected considerably less often during the second test session. It is difficult to rationalise this fully and future (larger cohort) test-retest studies using the ADI will need to carefully assess whether this is a consistent finding. The sensation of low intensity warmth is associated with activation of TRV3 or TRPV4 (Jordt et al., 2003), yet these channels are not triggered by menthol, nor does menthol have a heating effect on the skin (Wasner et al., 2004). Sensation of heat is associated with activation of TRP heat channels. It is unlikely that TRPV1 was activated since c-fibres expressing TRPV1 co-express TRPA1 and additional pronociceptive peptides but not TRPM8 (although this is still a matter of debate). TRPV1 / TRPA1 channel activation is also associated with a more intense burning or stinging sensation (Namer et al., 2005). In this study, some individuals did report burning sensation but this remained stable whilst warm / hot sensations changed between test occasions, suggesting distinct mechanisms. Basic science studies have reported that some low threshold nociceptors (Aδ and c2-fibres) (Campero & Bostock, 2010) are activated both by heat and cold. Hypothetically, temporary changes in spinal disinhibition might also explain the reduction in warm / heat experienced on the second occasion. Under normal circumstances cold signals (via non-nociceptive Aδ fibres) inhibit c-fibre signals in the dorsal horn. It may be that experience of heat on the first application of menthol was a reflection of a temporary central disinhibition of c-fibres, potentially driven supra-spinally by initial anxiety about an unknown experience. Once the application of menthol had been found to be manageable on the first occasion, the second test occasion for these individuals followed a more normal
cold-dominant pathway. This hypothesis is supported in part by a very recent study by
Thibodeau et al. (2012), which suggests that pain-related fear can increase
hyperalgesia during QST. It is an interesting speculation, which may deserve future
investigation.

**VAS reliability**

Although overall reliability was lower for VAS cold, heat, unpleasantness and pain it
was still at a relatively high level (ICCs .886 to .931). However, variability may have
been more of an issue for VAS scores: 95% confidence intervals showed greater
variability, particularly for heat and unpleasantness and SEM values were also largest
for these two sensations (Table 6.3). Scatterplots for cold, heat and unpleasantness
showed similar good levels of association between days (Figures 6.6a and b) and,
although pain VAS was poorly correlated between days, this reflects the high numbers
of participants registering no pain. Figure 6.5b illustrates that there was a clear
reduction in all VAS values by Day 2, most marked for unpleasantness. Given that the
menthol concentration and formulation, the test environment and the rater were
exactly the same between test sessions, it seems most likely that intensity levels were
also influenced by initial apprehension about an unknown sensory experience.
Intensity reduction on Day 2 is unlikely to have resulted from reduced activation of
receptors by menthol since the parameters were the same as for Day 1. Initial
apprehension may have translated into a temporary augmentation (or loss of normal
inhibition) of signals at central terminals in the dorsal horn, experienced as greater
intensity. It is significant that the percentage of subjects reporting any VAS for each
sensation remained relatively stable from one test occasion to the next (Figure 6.5a).
This indicates that, even though there may be some variation in intensity levels, the
nature of the sensation does not change significantly from one occasion to the next: if
an individual did not feel unpleasantness on the first occasion, they were unlikely to
report it on the second occasion.

**Time-course reliability**

An interesting finding from this study was for time-course of VAS responses. Despite
definitely visible similarities in the pattern of response over the 15-minute application at
Day 1 and Day 2 (Figures 6.8a), time to onset of first VAS sensation showed less
reliability than anticipated, particularly for cold intensity ICC .675). On Day 1 cold VAS
was first reported on average at two minutes into the application and about one minute
later on Day 2. This may reflect variability in skin permeation, which was not assessed
in the current investigation and will need to be reviewed in future studies with menthol gel. An alternative explanation may relate to initial apprehension once again, perhaps causing some hypervigilance to sensory change. Although this indicates that test application needs to ensure a standardised approach with individuals (either to concentrate fully on the sensation or to only take note of significant sensations), it is important to note that the degree of variability, even with cold VAS, was very small. Since the test aims to differentiate between relatively extreme responses, this level of variability may not be practically significant. The low overall variability in total ADI score also indicates that the weighting system enables the impact of variations in VAS intensity to be minimised.

In contrast to time to onset, time to peak showed greater reliability (ICCs .858 to .964) and low variability (Figure 6.6b). This may be a more true reflection of the reliability of menthol formulation release. The times to peak correspond well with those reported in Study 2 (Chapter 4) and also with the previous few studies which have applied menthol as a sensitiser for thermal cold sensation. All have reported peak sensation occurring at between 8 and 10 minutes for most individuals.

**Minimal gender differences**

This study briefly considered whether gender may play a part in any variability in scores. In support of previous studies looking at test-retest reliability of CPT, there were really no consistent differences between males and females on either test occasion. There was a non-significant trend towards increased report of heat and unpleasantness by males, and of increased pain by females. It is possible that these results point to a gender difference in word preference and further investigation into this may be valuable.
6.7 Summary and Conclusions

This study found that the menthol ADI test exhibited high levels of reliability when evaluated in healthy subjects over two test occasions. Although there was some variability in intensity values, with VAS ratings at Day 2, 5-8% lower than at Day 1, the quality if sensation did not appear to change.

A note of caution must be advised with these results so that only limited conclusions can currently be drawn. This is the first test-retest study for any menthol formulation and the first for the ADI scoring system. Further larger studies are clearly needed to review whether the same differences in intensity for the second test session are found and how much this impacts on score accuracy. One of the difficulties with topical menthol is that a practice test is impractical. Very mild menthol sensory effects can linger up to an hour after menthol removal in some people. Even applying the menthol to a different site the individual may be distracted by the ongoing sensation from the first site, so adding an additional confounding variable to the test. Consequently it may be that statistical analysis of a single occasion test may need to take a 5-10% learning effect into account. In addition, it must be acknowledged that, although this study showed good short-term test-retest reliability, it is not yet known if the test is reliable over longer periods of time. More data are needed to clarify both of these issues.

---

2 Appendix 3 provides basic data from a post-hoc test-retest analysis of Placebo group subjects from OA Study 7 (Chapter 9) in order to evaluate whether the current study reliability and variability findings were repeated. This analysis calculated ICCs for menthol test ADI score at baseline (Day 0) and on Day 4 for subjects who were drug-free on both test occasions. The OA-Placebo results showed similar levels of reliability for ADI (ICC=.917) and MWS (ICC=.886). Reliability for VAS cold, unpleasantness and pain were also comparable: ICCs: cold .889; unpleasantness .873; pain .897. However reliability for heat was considerably lower: ICC=.674. All VAS values in the Placebo study increased by Day 2, with heat intensity increasing by more than 60%. The implications of this are briefly considered in Appendix 3.
Chapter 7

Study Five

Matched comparison between healthy subjects and those with knee osteoarthritis

7.1 Abstract

Background and Aims
It has been proposed that cold hyperalgesia is associated with widespread thermal and mechanical hyperalgesia. Cold hyperalgesia may also be associated with increased pain and dysfunction and shows a higher prevalence in patient cohorts. This study compared the efficacy of the menthol test and conventional CPT in identifying cold hyperalgesia in individuals with and without pain. Differences in mechanical and thermal pain thresholds and physical function were analysed between subjects with painful knee OA and matched healthy controls.

Method
Forty healthy volunteers and 40 age and gender matched subjects with mild to moderately painful knee OA were recruited (16 male; 24 female, mean age 64 years for both groups). Self-reported pain and dysfunction were assessed with WOMAC and quality of life with the SF36 questionnaire. Physical function was tested with the timed Aggregated Locomotion Function (ALF) score. QST was assessed in randomised order: the menthol test (as used in Chapter 4, Gel B) was applied at the forearm and mechanical and thermal hyperalgesia tested at both knees and one elbow. Standard equipment, method and instructions were used.

Results
There were good correlations between ADI score and all CPT values (r = .548 to r = .605, p < .001). Incidence of cold hyperalgesia was significantly higher in the OA group whether assessed with CPT or ADI values: CPT > 15°C: 43% OA subjects versus 5% controls; ADI ≥ 5: 33% OA versus 15% controls. ADI scores were significantly higher in OA subjects (p = .002) with higher VAS unpleasantness and pain intensities and a higher percentage selecting unpleasant words: e.g. burning selected by 35% OA versus 10% controls. PPT and PPT values were higher for OA subjects at all sites but only reached significance at the OA knee (p < .001 and p = .020). Heat pain threshold was not significantly different at any site. OA subjects completed each ALF task significantly slower (p < .001 each task).
Conclusion
Subjects with painful knee OA showed widespread cold hyperalgesia, localised mechanical and punctate hyperalgesia and significantly reduced function compared with age and gender matched controls. The similarities between menthol ADI and CPT suggest that the menthol test is an equally valid instrument with which to identify individuals with cold hyperalgesia.
7.2 Introduction & Background

Studies 1-4 demonstrated the validity and reliability of the menthol test and Algoteck Descriptor Index (ADI) measurement system as an outcome measure for cold hyperalgesia in healthy subjects. Study 3 (Chapter 5) demonstrated that the ADI was able to predict membership of the CPT >15°C group with a sensitivity of 90% and a specificity of 74%. Study 4 (Chapter 6) showed that the ADI demonstrated good reliability between two repeated test sessions. However, the clinical value of a test for cold hyperalgesia lies in the accuracy with which it is able to identify individuals who report more severe chronic pain and functional disability.

There is increasing evidence across a range of conditions to suggest that patients who report higher levels of pain and dysfunction and poor response to standard interventions may also exhibit augmented pain processing. This has been demonstrated across conditions as wide-ranging as irritable bowel syndrome (Verne et al., 2012) endometriosis (Bajaj, 2003), carpal tunnel syndrome (Zanette et al., 2010) and temporal-mandibular joint disorder (Fernandez-de-las-Penas et al., 2009). In the rheumatology field, studies in fibromyalgia, rheumatoid arthritis and osteoarthritis have shown a close association between severity of pain and disability and extent of ‘central sensitisation’ (Goldenberg et al., 2011; Lee et al., 2011b; Murphy et al., 2012). In these studies, augmented pain processing, or ‘centralised pain’, is defined by a heightened response that spreads well beyond the usual anatomical location of the injured tissues. In experimental studies, this may also exhibit as abnormal wind-up or prolonged pain after cessation of input (Staud et al., 2007).

Many studies in osteoarthritis have used quantitative sensory testing methods to demonstrate widespread hyperalgesia. Kosek and Ordeberg (2000) reported lower pressure pain thresholds (PPT) in both the affected and contralateral hips in patients awaiting total hip replacement. Imamura et al. (2008) found that subjects with knee OA exhibited lower PPTs compared with controls, at sites as distant as S1 and quadratus lumborum. Widespread mechanical hyperalgesia correlated with higher pain intensity, higher disability scores, and with poorer quality of life. Infusion of hypertonic saline into tibialis anterior has been found to cause increased hyperalgesia over a wider area in people with knee OA compared with controls (Bajaj et al., 2001). Bradley et al. (2004) reported increased VAS pain response to pressure stimuli at the shoulder in subjects with knee OA. fMRI studies have reported increased punctate hyperalgesia in pain
referral areas associated with increased periaqueductal gray activation in subjects awaiting hip joint replacement (Gwilym et al., 2009).

Cold pain thresholds have also been found to be significantly elevated and associated with increased pain and disability in subjects with chronic pain. Park et al. (2010), found increased cold pain thresholds (CPT) in patients with arthrogenic temporomandibular joint disorder. Sterling et al. (2006), found that increased cold pain thresholds assessed within a month of whiplash injury (mean 18.7°C) were predictive of disability at two years post injury (mean 18.2°C). Bilateral cold hyperalgesia (22-24°C) in the hands of subjects with unilateral carpal tunnel syndrome was found to be associated with higher pain intensity and longer condition duration (de la Llave-Rincon et al., 2009). There is less data available for cold pain thresholds in osteoarthritis. Wylde et al. (2010) reported no difference in CPT between patients awaiting total knee replacement and healthy controls. In contrast, our own previous studies have consistently found significantly elevated CPT in subjects with hip (Wright et al., 2010)\(^1\) and knee OA (Moss et al., 2008)\(^1\) when compared with matched healthy controls. However there is still limited evidence for the association between cold hyperalgesia and report of high pain severity and dysfunction in people with knee OA.

This study therefore assessed the incidence of cold hyperalgesia in individuals with painful knee OA, compared with an equivalent matched healthy control group, using both conventional thermode-assessed CPT and the menthol test to evaluate response to cold. Additional QST data for mechanical, punctate and heat pain thresholds at local knee and distant upper limb sites was collected to explore associations between cold and other QST modalities. A series of everyday activity tests was used to examine associations between hyperalgesia and physical function.

Response to the menthol test in older and clinical cohorts was examined in this study. It has been suggested that older individuals may report pain as less intense and using different quality descriptors (Gagliese & Katz, 2003; Gagliese & Melzack, 2003). It was therefore necessary to compare responses to menthol of healthy older adults (aged 50 years and over) in this study with that of younger subjects from earlier studies. In addition, this study was able to compare the specific intensity and quality responses of subjects with a chronic pain disorder against those of their healthy counterparts.

---

\(^1\) Refer Appendix 5
The primary aim of this study therefore was to assess the efficacy of the menthol test in identifying cold hyperalgesia in older individuals with chronic pain from knee osteoarthritis compared with age and gender matched pain-free controls. The study also explored whether mechanical, punctate and heat hyperalgesia and physical function were also reduced in those with chronic pain compared with matched controls.

### 7.3 Hypotheses:

1. The intensity, quality and release response characteristics to menthol gel for healthy older subjects will be consistent with that of previous studies.

2. There will be a significant difference in response to the menthol test between subjects with knee OA and healthy matched controls.

3. There will be good associations between cold hyperalgesia as measured with ADI and cold hyperalgesia as measured with CPT.

4. There will be significant differences in pain thresholds for cold, heat, pressure and punctate stimuli but no significant differences in sensory detection thresholds between subjects with knee OA and healthy matched controls.

5. There will be significant differences in physical function and quality of life between subjects with knee OA and healthy matched controls.
7.4 Methods

Subjects

Forty subjects with painful knee osteoarthritis and 40 matched pain-free healthy controls were voluntarily recruited from the Perth community via radio and newspaper advertisements. Volunteers with painful knee OA (WOMAC subscale score ≥3) were recruited and assessed for suitability by initial phone screen and then by a Rheumatologist, using the clinical criteria listed below (Table 7.1).

<table>
<thead>
<tr>
<th>Inclusion:</th>
<th>Aged 50 years +</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unilateral diagnosis of knee OA &gt; 6 months [American College of Rheumatology clinical classification system];</td>
</tr>
<tr>
<td></td>
<td>Pain in the index knee &gt;3/10 on a pain visual analogue scale;</td>
</tr>
<tr>
<td></td>
<td>No additional clinically significant joint involvement;</td>
</tr>
<tr>
<td></td>
<td>In good health other than OA knee;</td>
</tr>
<tr>
<td></td>
<td>No arthroscopy or injections to index knee in last 6 months.</td>
</tr>
<tr>
<td></td>
<td>Able to read and understand English</td>
</tr>
</tbody>
</table>

| Exclusion: | History of systemic inflammatory conditions; |
|           | Neurological deficits (motor, cognitive or sensory); |
|           | Recent lower limb injury or surgery (< 6 months); |
|           | History of other chronic pain disorders (eg fibromyalgia); |
|           | Skin allergies; |
|           | Allergy to menthol. |

Table 7.1: Inclusion / Exclusion criteria for subjects with knee OA

Healthy volunteers were assessed for suitability via phone screen and matched by gender and 10-year age band to the 40 OA subjects using inclusion / exclusion criteria below (Tables 7.2 and 7.3).
Table 7.2: Inclusion / exclusion criteria for healthy control subjects

<table>
<thead>
<tr>
<th>Inclusion:</th>
<th>Aged 50 years +</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In general good health, no current pain.</td>
</tr>
<tr>
<td>Exclusion:</td>
<td>Diagnosis of lower limb OA;</td>
</tr>
<tr>
<td></td>
<td>History of systemic inflammatory conditions;</td>
</tr>
<tr>
<td></td>
<td>Neurological deficits (motor, cognitive or sensory);</td>
</tr>
<tr>
<td></td>
<td>Recent lower limb injury or surgery (&lt; 6 months);</td>
</tr>
<tr>
<td></td>
<td>History of other chronic pain disorders (e.g. fibromyalgia);</td>
</tr>
<tr>
<td></td>
<td>Skin allergies.</td>
</tr>
<tr>
<td></td>
<td>Allergy to menthol</td>
</tr>
</tbody>
</table>

Table 7.3: Subjects with knee OA and controls were matched by 10-year age band and gender.

<table>
<thead>
<tr>
<th>Gender (male : female)</th>
<th>OA (n=40)</th>
<th>Control (n=40)</th>
<th>t(78)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age: mean (range) years</td>
<td>64 (50-80)</td>
<td>64 (50-81)</td>
<td>.121</td>
<td>.904</td>
</tr>
<tr>
<td>Age spread (n)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>16:24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>16:24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-79</td>
<td>16:24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80-89</td>
<td>16:24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All participants provided written informed consent before participating in the study. Ethical approval was provided in a single submission for studies 5, 6 and 7 by Royal Perth Hospital Medical Research Ethics Committee (Approval EC 2009/100) and by Curtin University Human Research Ethics Committee (Approval number HR 26/2010).

**Study Design and Procedures**

The study used a cross-sectional design, with subjects attending for testing on one occasion, at which all outcome measures were assessed. Following brief subjective examination to ensure inclusion and exclusion criteria had been satisfied and completion of informed consent procedures, subjects completed the SF-36 (short form) questionnaire. This quality of life survey is widely accepted as the ‘gold standard’ across a wide range of physical disorders (Angst et al., 2001; Gandhi et al., 2001). OA subjects also completed the Western Ontario and McMaster Universities Osteoarthritis Index for the Knee (WOMAC) (Bellamy, 1988). All subjects then completed physical and QST outcome measures. Standardised method and instructions were used for each test. Order of testing was randomised between modalities, although for punctate, heat and cold stimuli, detection threshold was always tested before pain threshold.
**Outcome Measures**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>QST</td>
<td>• <em>Algotect Descriptor Index (ADI):</em></td>
</tr>
<tr>
<td></td>
<td>o VAS intensity for cold, heat, unpleasantness and pain</td>
</tr>
<tr>
<td></td>
<td>o Pain quality (MWS)</td>
</tr>
<tr>
<td>Cold detection</td>
<td>• <em>Cold detection threshold (CDT) (°C)</em></td>
</tr>
<tr>
<td>Thermal cold hyperalgesia</td>
<td>• <em>Cold pain threshold (CPT) (°C)</em></td>
</tr>
<tr>
<td>Heat detection</td>
<td>• <em>Heat detection threshold (HDT) (°C)</em></td>
</tr>
<tr>
<td>Heat hyperalgesia</td>
<td>• <em>Heat pain threshold (HPT) (°C)</em></td>
</tr>
<tr>
<td>Light touch detection</td>
<td>• <em>Punctate detection threshold (PcDT)</em></td>
</tr>
<tr>
<td>Light touch allodynia</td>
<td>• <em>Punctate pain threshold (PcPT)</em></td>
</tr>
<tr>
<td>Mechanical hyperalgesia</td>
<td>• <em>Pressure pain threshold (PPT) (kPa)</em></td>
</tr>
<tr>
<td>Physical Function</td>
<td><strong>Timed (secs):</strong></td>
</tr>
<tr>
<td>Aggregated locomotion function score</td>
<td>• 2m return chair transfer</td>
</tr>
<tr>
<td>ALF (Aggregated Locomotion Function)</td>
<td>• 8m return walk</td>
</tr>
<tr>
<td></td>
<td>• Ascent/descent 10 stairs</td>
</tr>
<tr>
<td>Self-Report</td>
<td><strong>SF36-v2</strong></td>
</tr>
<tr>
<td>Quality of Life</td>
<td><em>Physical Health Score</em></td>
</tr>
<tr>
<td></td>
<td><em>Mental Health Score</em></td>
</tr>
<tr>
<td>Pain and function</td>
<td>• <em>Western Ontario &amp; Macmaster Universities (WOMAC) Osteoarthritis Knee Index</em></td>
</tr>
</tbody>
</table>

**Quantitative Sensory Tests:**

All quantitative sensory tests were applied at standardised sites. Menthol-cold hyperalgesia was tested at the standardised site on the volar forearm, 10cms above the wrist crease. Other QST tests were applied at the OA (or control equivalent) knee, the uninvolved knee and the ipsilateral elbow (Table 7.4). OA and “unaffected” knees were designated according to highest VAS pain rating.

**Table 7.4:** Standardised knee and elbow test sites for QST tests

<table>
<thead>
<tr>
<th>Punctate &amp; Pressure Thresholds</th>
<th>Thermal Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knee</td>
<td>Medial knee joint line</td>
</tr>
<tr>
<td>Elbow</td>
<td>Common extensor origin</td>
</tr>
</tbody>
</table>
1. **Menthol Test.** This study used menthol Gel B. Response was measured using the e-VAS and descriptors described in Appendix 2. A single batch was prepared for use with all control and OA subjects in Studies 5 to 7. Once the test area skin on the volar forearm had been cleaned with hypoallergenic soap and lukewarm water, 2mls of menthol gel was applied using a small syringe. A Tegaderm dressing was immediately placed over the gel, which was gently spread out to the area of the 2x3cm window. The electronic VAS timer was immediately started for the 15-minute application. At one minute intervals the response box light alerted subjects to rate the intensity of cold, heat, unpleasantness and pain they were experiencing by moving the four 100mm sliders. The ends of each slider were marked with ‘maximum’ and ‘minimum’ for one of the four sensations. Further details of the e-VAS system can be found in Appendix 2. Every two minutes subjects were asked to select as many or few words as they wished from the Descriptor List. After 15 minutes, the Tegaderm and gel were removed using hypoallergenic soap and tepid water, gently patted dry with a paper towel. Algotect Descriptor Index (ADI) was calculated. In addition, values were calculated for maximum VAS intensity for each sensation, time to onset and peak for each VAS sensation. From the descriptor list, the Mean Word Score (MWS) was also recorded.

2. **Cold Detection and Cold Pain Thresholds (CDT & CPT)** were measured using a Peltier thermode (Medoc, Israel) and standard Method of Limits (Rolke, 2006; Heldestad et al., 2010). The 3x3cm contact probe was attached to the test site with a velcro strap and subjects given several minutes to adapt to the baseline temperature of 32°C. The thermode temperature was set to drop at a constant rate of 1°C/sec to a minimum of 0°C. For both cold detection and cold pain tests, standardised positioning and instructions were used. Cold detection threshold was always measured first in order to avoid desensitisation from pain threshold testing. Subjects were instructed to depress the hand-held switch as soon as they perceived any cooling change from baseline. The temperature (°C) was recorded, with the thermode returning to baseline at 1°C/sec. For cold pain threshold, subjects were instructed to press the switch as soon as the cooling sensation changed to one of painful cold. This temperature was recorded (°C) and the thermode returned to baseline again. For both CDT and CPT at each test site, one practice was followed by 3 trials, each trial separated by a randomly assigned pause of between 3 and 6 seconds, as for detection thresholds. The mean of the trials was calculated for
3. **Heat Detection and Pain Thresholds (HDT & HPT)** were also measured with the Medoc Peltier thermode using similar methodology as for CDT / CPT (baseline 32°C, 1°C/sec ascending and descending ramps), with maximum temperature set at 50°C. Standardised positioning and instructions were again used. Heat detection was tested before heat pain threshold for similar reasons as for cold testing. HDT was defined as the temperature (°C) at which subjects first perceived an increase in warmth from baseline, whilst HPT was defined as the temperature (°C) at which subjects perceived that the heating sensation had become one of painful heat. As for CDT / CPT, as soon as the subject pressed the switch to signal threshold, the temperature returned to baseline. For both HDT and HPT, trials were again separated by randomly assigned variable pauses of between 3 and 6 seconds. One practice was followed by 3 trials, with the mean calculated for analysis.

4. **Punctate Pressure Detection and Pain Thresholds (PcDT & PcPT)** were assessed using von Frey filaments (Semmes Weinstein monofilaments, Bioseb) at each of the standardised test sites. For both detection and pain thresholds, the staircase method was applied, applying probe tips in ascending and then descending order of magnitude for three repetitions. Punctate detection threshold (PcDT) was defined as the smallest von Frey tip (g) perceived as touch sensation, whilst punctate detection threshold was defined as the smallest tip (g) perceived (Rolke, 2006). Punctate pain threshold was the smallest von Frey tip felt as a “pinprick” sensation rather than just light touch. A practice plus three readings at a standardised position at each joint was taken for each measure, with the geometric mean used for analysis.

**Physical Function Testing**
The Aggregated Locomotor Function (ALF) score (McCarthy & Oldham, 2001) was used as a measure of observed locomotor function. The score was calculated by summing the timed values (seconds) for 3 locomotor tasks: time taken to walk 2-metres to a chair, sit, stand and walk back 2-metres; time taken to complete an 8-metre return walk; and ascent / descent of 10 stairs. All instructions were carefully standardised, with subjects asked to complete each task “as briskly as possible”. A stopwatch was used for timing, as advised in the test instructions. This score has shown excellent inter-rater reliability and low standard error of measurement (McCarthy & Oldham, 2004). The score is
moderately well correlated with both WOMAC and SF-36 function indices, and is reported to be responsive to change following intervention over a short time period (McCarthy & Oldham, 2004)

**Self-Report Questionnaires**

1. *SF-36v2™ Health Survey*. Quality of Life for knee OA and control subjects was measured with the SF-36v2, a generic self-reported health outcomes assessment tool. The tool measures the self-perceived impact of health status on quality of life via eight domains, using Likert-type response categories. The SF-36v2 electronic scoring system was applied to generate the physical and mental health sub-scores for analysis (range 0-100 for each sub-scale). All versions of the SF-36v2 have demonstrated validity and reliability when used with a range of conditions (Gandhi et al., 2001).

2. *Western Ontario and McMaster Universities Osteoarthritis Index for the Knee (WOMAC)* was used to evaluate baseline subjective pain, stiffness and functional limitation for OA subjects only. This OA-specific self-report scale has been widely used to measure pain and disability from knee OA, demonstrating good internal validity and test-retest reliability (Jinks et al., 2002). WOMAC was collected with specific reference to the index knee in this study. The visual analogue scale version was used, with responses to each of the 24 questions marked on a 10cm line. Scores could range from 0-50 for the pain subscale, 0-20 for stiffness and 0-170 for function. Total WOMAC score therefore ranged from 0-240.

**Sample Size and Statistical Analysis**

Since no previous studies have compared response to menthol between healthy subjects and those with pathology, a sample size of 40 per group was calculated using previous QST data. Previous studies have shown that a sample size of between 37 and 45 would be needed to detect a clinically significant difference of 15% in PPT and CPT between healthy and OA groups, with a power of 80% and significance level of 0.05. This amounts to a between groups actual difference of 38kPa and mean standard deviation 57kPa for PPT (Moss et al., 2007) and a between groups difference of 5.4°C, mean standard deviation of 2.3°C for CPT (Moss et al., 2008).

---

2 All permissions granted for use of both questionnaires from copyright owners
Chapter 7  
Comparison OA: Normal

Normality testing was carried out using a combination of Shapiro-Wilk tests and visual analysis of distribution. ADI, MWS, cold VAS, PPT, ALF and WOMAC data were normally distributed (Shapiro-Wilk p = .064 to .307) so that parametric statistics could be applied. Unpleasantness and pain VAS data required non-parametric analysis. CPT data showed its usual bimodal distribution when reviewed as a single dataset, however was normally distributed once divided into CPT < or >15°C subsets. Accordingly, when CPT was analysed as a single dataset, non-parametric analysis was applied, but when data was divided according to CPT < or >15°C, parametrics were applied.

Data were analysed using the SPSS statistical package, version 19. Alpha was set at p<.05. The following analyses were carried out:

<table>
<thead>
<tr>
<th>Hypotheses and Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The intensity, quality and release response characteristics to menthol gel for healthy older subjects will be similar to that for healthy subjects in Study 2:</td>
<td></td>
</tr>
<tr>
<td>• Maximum VAS rating</td>
<td>Descriptive comparisons between studies</td>
</tr>
<tr>
<td>• Time to 1st sensation (mins)</td>
<td></td>
</tr>
<tr>
<td>• Time to peak sensation (mins)</td>
<td></td>
</tr>
<tr>
<td>• MWS and ADI values</td>
<td>Spearman’s Correlation Coefficient</td>
</tr>
<tr>
<td>2. Difference in response to menthol between subjects with knee OA and healthy matched controls:</td>
<td></td>
</tr>
<tr>
<td>• Maximum VAS rating</td>
<td></td>
</tr>
<tr>
<td>• Time to 1st sensation (mins)</td>
<td>Mann-Whitney U tests</td>
</tr>
<tr>
<td>• Time to peak sensation (mins)</td>
<td></td>
</tr>
<tr>
<td>• MWS and ADI values</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>3. Differences in pain thresholds for cold, heat, pressure and punctate stimuli:</td>
<td></td>
</tr>
<tr>
<td>• CPT (°C)</td>
<td>Mann-Whitney U tests</td>
</tr>
<tr>
<td>• HPT (°C)</td>
<td>Or</td>
</tr>
<tr>
<td>• PPT (kPa)</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>• PcPT (log value)</td>
<td></td>
</tr>
<tr>
<td>4. Differences in detection thresholds for cold, heat and punctate stimuli:</td>
<td></td>
</tr>
<tr>
<td>• CDT (°C)</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>• HDT (°C)</td>
<td></td>
</tr>
<tr>
<td>• PcDT (log value)</td>
<td></td>
</tr>
</tbody>
</table>
5. Differences in measures of physical function and quality of life:
   - *ALF: chair, walk, stairs tasks (secs)*
   - *SF36: PCS and MCS values*

   **Independent t-tests**

6. Associations between measures of response to cold (CPT and ADI) and additional measures of function (ALF) and hyperalgesia (PPT).

   **Pearson’s or Spearman’s Correlation Coefficients**
7.5 Results

Characteristics of the healthy and OA cohorts

Although gender and age were very closely matched (Table 7.3), there was a significant difference in BMI between groups, with OA subjects having a significantly higher mean BMI than their healthy counterparts (p=.001) (Table 7.5). SF-36 scores showed that OA subjects reported significantly reduced quality of life due to physical function compared with control subjects (p=.046) however there was no significant group difference in perceived quality of life due to mental health issues (p=.961). OA subjects also completed the disease-specific questionnaire WOMAC. This cohort of 40 subjects with knee OA reported a mean pain level of 14.5/50 (SD 7.8), mean stiffness of 7/20 (SD 4.1) and functional level of 49.7/170 (SD 27). Mean pain, stiffness and disability was therefore only mild to moderate (30-35% of highest possible scores).

Table 7.5: Comparisons between OA and matched control subjects for BMI and quality of life.

<table>
<thead>
<tr>
<th></th>
<th>OA (n=40)</th>
<th>Control (n=40)</th>
<th>t(78)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI mean (SD)</td>
<td>28.4 (7.7)</td>
<td>25.3 (3.2)</td>
<td>3.56</td>
<td>.001**</td>
</tr>
<tr>
<td>BMI band (n)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight (&lt;19)</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal (19-24)</td>
<td>7</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight (25-30)</td>
<td>20</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese (&gt;30)</td>
<td>13</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(\chi^2 = 8.96)</td>
<td>.011*</td>
</tr>
<tr>
<td>SF-36 Quality of Life mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical sub-score</td>
<td>39.3 (7.7)</td>
<td>43.6 (10.0)</td>
<td>-2.03</td>
<td>.046*</td>
</tr>
<tr>
<td>Mental health sub-score</td>
<td>54.2 (7.1)</td>
<td>54.3 (10.6)</td>
<td>-.049</td>
<td>.961</td>
</tr>
</tbody>
</table>

Gender Differences

Across all subjects, there was no significant difference between genders for CPT or ADI values, although PPT at the OA knee was significantly higher for females.

There was no significant difference between males and females for CPT or ADI values (CPT at the OA knee \(t_{(78)} = -1.04, p=.281\); ADI \(t_{(78)} = -0.577, p=.565\). There was however a gender difference in PPT, with females exhibiting significantly lower PPT values at the OA knee \(t_{(78)} = 2.72, p=.009\).
Distribution of ADI Data

**Figure 7.1:** Histogram showing the normal distribution curve for menthol ADI data across all subjects. The Shapiro-Wilk test (p = .223) suggested that the data was from a normally distributed sample.

**Hypothesis 1:**

The intensity, quality and release response characteristics to menthol gel for older healthy control subjects will be consistent with that of previous studies.

- Characteristics of healthy subjects in Study 2 (Chapter 4) compared with those in the current study.

Healthy subjects in Study 2 and the current study were all pain-free and had no history of chronic pain disorder. There was a similar gender balance between the healthy subjects of Study 2 and those of the current study: Study 2 comprised 40% male and 60% female; the current study 37% male and 63% female. Mean and range of age however was significantly higher in the current study: mean 64 years, range 50-81 in comparison to Study 2 mean 34 years, range 18-52 years.

- Sensation intensity was lower and unpleasant qualities reported by fewer older healthy control subjects

90% of control subjects in this study reported some cold intensity during the 15-minute menthol application. This is similar to the 94% reported for healthy subjects in Study 2. However, a lower percentage of healthy subjects in this study reported heat, unpleasantness and pain compared with Study 2 (heat: 17% versus 52% for Study 2; unpleasantness 30% versus 74%; pain 11% versus 52%). Maximum VAS was also lower for the current study: cold - 24 versus 35/100; heat – 17 versus 24/100; unpleasantness 19 versus 24/100; pain 15 versus 18/100) (Figure 7.3b).

Older healthy subjects in the current study selected predominantly cold-type words and the mildly unpleasant dysesthetic word tingling to describe the sensation during
menthol application (Figure 7.6 below). More intense words such as icy/freeing, burning and stinging or were only selected by a minority of subjects (12.5%, 10%, and 20% respectively). Healthy subjects in Study 2 had a more intense response to the menthol gel, with 48% reporting icy/freeing, 41% burning and 26% stinging. Milder dysesthetic words were selected by similar percentages.

- **Release times were comparable between studies**

  VAS release characteristics and time course were generally comparable between this healthy cohort and those from Study 2 (Figure 7.2a and b). Cold was reported first at between 2.5 to 3 minutes after initial application start for both studies. Heat was experienced at around 5.5 minutes for both also. Unpleasantness was felt slightly earlier and pain almost 2 minutes earlier in Study 2. There was also a slight difference in time to peak intensity, with all sensations reported at around 8.5 minutes for Study 2, but with a greater spread for the current study, from 7.5 minutes for cold and heat to over 9 minutes for pain.

![Figure 7.2a](image1.png): Healthy older control subjects from current study

![Figure 7.2b](image2.png): Healthy subjects from Study 2

**Figure 7.2a-b:** Mean time to onset and time to peak for each VAS sensation.

- **Within the current study, there were good associations between intensity and quality values for healthy subjects**

  The current study found moderate to good positive significant correlations within the menthol test between descriptor sub-score (MWS) and cold, heat and unpleasantness
Chapter 7

Comparison OA: Normal

VAS intensities: cold r = .418, p = .007; heat r = .527, p < .001; unpleasantness r = .541, p < .001. There was no correlation between MWS and VAS pain: r = .177, p = .276.

- **Within the current study, there were also good associations between menthol test score and CPT temperature.**

The current study found good correlations between ADI score and CPT temperatures and even stronger correlations between MWS descriptor sub-score and CPT temperature (Table 7.6). All VAS intensity ratings also significantly correlated with CPT but only at a low level, ranging from r = .420, p < .001 for VAS cold at the elbow to r = .222, p = .048 at for VAS unpleasantness at the unaffected knee.

<table>
<thead>
<tr>
<th>Table 7.6: Correlations between ADI and MWS scores and CPT temperature at the OA knee, unaffected knee and elbow.</th>
</tr>
</thead>
<tbody>
<tr>
<td>**** ****</td>
</tr>
<tr>
<td><strong>OAD</strong></td>
</tr>
<tr>
<td><strong>r</strong></td>
</tr>
<tr>
<td><strong>p</strong></td>
</tr>
<tr>
<td>ADI</td>
</tr>
<tr>
<td>p &lt; .001**</td>
</tr>
<tr>
<td>MWS</td>
</tr>
<tr>
<td>p &lt; .001**</td>
</tr>
</tbody>
</table>

**Hypothesis 2:**

*There will be a significant difference in response to the menthol test between subjects with knee OA and healthy matched controls.*

- **ADI score was significantly higher in OA subjects**

There was a significant difference in total ADI scores between OA and healthy subjects: OA mean 4.0 (± .28), healthy mean 2.8 (± .27), t(78) = -3.26, p = .002.

- **Sensation intensity was significantly greater for cold and heat for OA subjects**

A similar proportion of subjects in both OA and control groups reported some cold intensity during the menthol application. Heat was also reported by similar percentages of subjects. However, more than 50% of subjects with knee OA reported some unpleasantness, and 20% reported some pain, compared with only 30% and 11% of controls respectively (Figure 7.3a-b). There were significant differences between OA and control groups in maximum VAS rating. These differences reached statistical
significance for VAS cold ($t_{(78)} = -2.47, p = .016$), and unpleasantness ($t_{(78)} = -2.75, p = .007$), but not quite for pain ($t_{(78)} = -1.82, p = .075$) or heat ($t_{(78)} = -1.48, p = .142$).

**Figure 7.3a**: Subjects with knee OA.  
**Figure 7.3b**: Healthy control subjects

**Figure 7.3a-b**: Percentage of subjects reporting some VAS intensity (>0/100) for each sensation (bar chart, axis 1); Average maximum intensity (0-100) for those reporting VAS>0 for each sensation (line graph, axis 2)

• *Release: there was little difference in time to onset or peak between groups although intensity time course showed a different pattern (Figure 7.4).*

There were no statistically significant differences between OA and control groups in time to onset for any VAS sensation. Time to peak sensation was very similar for cold and heat, all peaking at between 7 and 7.5 minutes. Unpleasantness and pain peaked at between 8 and 9 minutes for both.

**Figure 7.4**: OA subjects: mean time to onset and time to peak for each VAS sensation

Visually, the time course for each sensation followed a slightly different pattern (Figure 7.5a-b). In particular, for subjects with OA, heat, unpleasantness and pain all rose throughout the menthol application, whereas for controls, intensity of these unpleasant sensations had started to decline well before menthol removal.
**Sensation quality:** both MWS and ADI scores were significantly higher for OA subjects

Mean descriptor MWS score was significantly higher for OA than control groups ($t_{(79)} = -3.28, p=.002$). This was reflected in the higher percentage of OA subjects who selected more intense words such as icy/freeing or more unpleasant words such as burning, prickling or penetrating. In contrast, healthy controls were more likely to select less intense words such as cool or tingling. Stinging was selected by similar percentages of subjects from both groups (Figure 7.6).

**Figure 7.5a-b:** Mean VAS intensity at each minute for each sensation

**Figure 7.5a:** Healthy control subjects

**Figure 7.5b:** OA knee subjects

**Figure 7.6:** Percentage choice of key descriptor words by healthy and OA subjects.
When weighted VAS intensity and word choice were combined into the total ADI a significant difference in values was seen between OA and control groups ($t_{[79]} = -3.26$, $p=.002$).

**Hypothesis 3:**

*There will be good associations between cold hyperalgesia as measured with ADI and cold hyperalgesia as measured with CPT.*

- *There were good correlations between ADI score and CPT temperatures across the whole cohort.*

ADI score showed good correlations with mean CPT ($r = .585$, $p < .001$) and also with CPT at each test site: OA knee $r = .548$, $p < .001$; unaffected knee $r = .605$, $p < .001$; elbow $r = .600$, $p < .001$).

- *More OA subjects were categorised as cold hyperalgesic using CPT > 15°C than using ADI ≥ 5.*

Subjects were divided dichotomously into menthol-cold hyperalgesic or not, using an ADI cut-off score of 5 and thermal-cold hyperalgesic or not using a mean CPT cut-off temperature of 15°C. This resulted in 13 (32.5%) subjects with knee OA categorised as menthol-cold hyperalgesic, in comparison to only 6 (15%) of healthy controls. For CPT values, 17 (42.5%) OA subjects were classified as cold hyperalgesic in contrast to only 2 (5%) healthy subjects. When membership of the cold hyperalgesic groups for ADI and CPT were cross-matched, 9 OA subjects (22.5%) were classified as cold hyperalgesic by both grouping methods (Table 7.7).

<table>
<thead>
<tr>
<th></th>
<th>ADI&lt;5</th>
<th>ADI ≥5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OA subjects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPT&lt;15°C</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>CPT &gt;15°C</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>27</td>
<td>13</td>
</tr>
<tr>
<td><strong>Healthy subjects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPT&lt;15°C</td>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>CPT &gt;15°C</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>34</td>
<td>6</td>
</tr>
</tbody>
</table>

*Table 7.7: Cross-tabulation of membership of high and low ADI and CPT groups by OA and healthy subjects.*
• Both classifications for cold hyperalgesia (ADI ≥ 5 and CPT > 15°C) showed similar ability to identify greater widespread mechanosensitivity and dysfunction. When cold hyperalgesic subjects were compared widespread mechanical hyperalgesia at all tests sites was consistently identified using both ADI ≥ 5 and CPT > 15°C cut-offs (Figure 7.7): ADI ≥ 5 cut-off (OA knee $t_{(70)}$ = 2.81, $p = .006$; unaffected knee $t_{(70)}$ = 2.44, $p = .017$; elbow $t_{(70)}$ = 2.57, $p = .003$); CPT > 15°C cut-off (OA knee $t_{(70)}$ = 2.76, $p = .007$; unaffected knee $t_{(70)}$ = 5.01, $p < .001$; elbow $t_{(70)}$ = 3.88, $p < .001$).

![Image](image.png)

**Figure 7.7:** Mean pressure pain thresholds at each test site, comparing subjects divided according to CPT < or > 15°C or ADI < or ≥ 5.

All cold hyperalgesic subjects were also slower at each of the ALF tasks, although significance was not reached for any grouping (Figure 7.8): ADI (chair transfer $t_{(70)}$ = -1.20, $p = .233$; walk $t_{(70)}$ = -1.97, $p = .053$; stairs $t_{(70)}$ = -1.08, $p = .284$); CPT (chair transfer $t_{(70)}$ = -1.15, $p = .253$; walk $t_{(70)}$ = -1.82, $p = .130$; stairs $t_{(70)}$ = -1.73, $p = .098$).

![Image](image.png)

**Figure 7.8:** Mean times for each ALF task, comparing subjects divided according to CPT < or > 15°C or ADI < or ≥ 5.
Hypothesis 4:

Subjects with knee OA and healthy matched controls will show significant differences in pain thresholds for cold, heat, pressure and punctate stimuli but no significant differences in sensory detection thresholds.

- Cold pain threshold was significantly higher at all sites for OA subjects. There was a significant difference between OA and control subjects in cold pain thresholds at the local OA affected knee (or equivalent) \( U_{78} = 370, z = -4.15, p < .001 \) and also at the unaffected knee \( U_{78} = 466, z = -3.23, p < .001 \) and at the distant elbow: \( U_{78} = 420, z = -4.39, p < .001 \). 43% of those with OA exhibited CPT >15°C at the OA knee and 30% exhibited CPT >15°C at the elbow, compared with only 7.5% and 5% of healthy control subjects respectively.

![Image of graph showing cold pain threshold comparison between OA and Healthy Controls]

**Figure 7.9:** Mean cold pain threshold at the OA knee, unaffected knee (or equivalents) and elbow for OA and healthy control subjects.

- Pressure pain threshold was significantly higher at the OA knee for OA subjects but not at other sites.

There was a significant group difference for PPT at the OA knee \( t_{78} = -3.68, p < .001 \). Significance was almost reached at the unaffected knee \( t_{78} = 1.93, p = .057 \) but there was no significant difference between groups at the elbow \( t_{78} = -.101, p = .919 \).

![Image of graph showing pressure pain threshold comparison between OA and Healthy Controls]

**Figure 7.10:** Mean pressure pain thresholds at the OA knee, unaffected knee (or equivalents) and elbow for OA and healthy control subjects.
• **Punctate pain threshold for OA subjects was also significantly lower at the OA knee but not at any other site.**

There was a significant group difference in PcPT at the OA knee ($t_{(78)} = -2.37, p=.020$). The difference did not reach significance at the unaffected knee ($t_{(78)} = -1.64, p=.104$) or at the elbow ($t_{(78)} = -.783, p=.463$).

• **Heat pain threshold was reduced for subjects with OA but the difference did not reach significance at any site.**

There was no significant difference between OA and control subjects for HPT at the OA (equivalent) knee ($t_{(78)} = 1.84, p=.070$), at the unaffected (equivalent) knee $t_{(78)} = .833, p=.407$) or at the elbow ($t_{(78)} = .956, p=.342$), although at each site, OA subjects were more sensitised to heat, with pain thresholds reached sooner than for controls.

![Fig 7.11: Mean heat pain threshold at the OA knee, unaffected knee (or equivalents) and elbow for OA and healthy control subjects](image)

• **OA subjects were less sensitive to cold detection at both knees, although only significantly at the OA knee. There were no differences between groups for heat or punctate detection thresholds.**

There was a significant group difference in CDT at the OA knee ($p = .104$) with OA subjects detecting cold at lower temperatures (Table 7.8). A similar although non-significant pattern was seen at the unaffected knee ($p = .071$). There was no significant difference at the elbow ($p = .259$). Heat and punctate detection thresholds were not significantly different at any site.
**Table 7.8:** Cold, heat and punctate detection thresholds for OA and healthy control subjects at each test site

<table>
<thead>
<tr>
<th></th>
<th>OA</th>
<th>Control</th>
<th>t(78)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cold Detection Threshold</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OA knee</td>
<td>28.8 (.22)</td>
<td>29.5 (.20)</td>
<td>-2.24</td>
<td>.028*</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>28.9 (.28)</td>
<td>29.5 (.18)</td>
<td>-1.83</td>
<td>.071</td>
</tr>
<tr>
<td>Elbow</td>
<td>29.3 (.30)</td>
<td>29.7 (.18)</td>
<td>1.14</td>
<td>.259</td>
</tr>
<tr>
<td><strong>Heat Detection Threshold</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OA knee</td>
<td>36.0 (1.4)</td>
<td>35.6 (1.6)</td>
<td>1.07</td>
<td>.290</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>36.2 (1.9)</td>
<td>35.7 (1.7)</td>
<td>1.31</td>
<td>.195</td>
</tr>
<tr>
<td>Elbow</td>
<td>35.8 (1.7)</td>
<td>35.4 (1.7)</td>
<td>1.11</td>
<td>.271</td>
</tr>
<tr>
<td><strong>Punctate Detection Threshold</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OA knee</td>
<td>3.7 (.04)</td>
<td>3.5 (.06)</td>
<td>1.62</td>
<td>.110</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>3.7 (.05)</td>
<td>3.7 (0.4)</td>
<td>.072</td>
<td>.943</td>
</tr>
<tr>
<td>Elbow</td>
<td>3.4 (.07)</td>
<td>3.1 (.07)</td>
<td>1.64</td>
<td>.105</td>
</tr>
</tbody>
</table>

**Hypothesis 5:**

*Subjects with knee OA and healthy matched controls will show significant differences in physical function and quality of life.*

- Subjects with OA took significantly more time to complete each of the ALF tasks

There was a significant difference between OA and healthy control subjects in time taken to complete the chair transfer ($t_{(78)} = 3.60$, $p = .001$), the 8m walk ($t_{(78)} = 4.71$, $p < .001$) and the stairs ($t_{(78)} = 34.62$, $p < .001$) (Figure 7.12).

![Figure 7.12](image-url)
• Subjects with OA reported significantly reduced quality of life due to physical difficulties but no difference in mental health-related quality of life.

There was a significant group difference in SF-36 physical sub-score (p=.046). There were no differences between mental health sub-scores (p=.961) (Table 7.9).

| Table 7.9: SF-36 physical and mental health sub-scores for OA and healthy control subjects |
|-----------------------------------------------|-------|-------|--------------|
| SF-36 Quality of Life                         | OA    | Control | t(78) | p      |
| Physical sub-score                           | 39.3 (7.7) | 43.6 (10.0) | -2.03 | .046* |
| Mental sub-score                             | 54.2 (7.1) | 54.3 (10.6) | -.049 | .961  |

Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Accepted/Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Response characteristics to menthol gel for healthy subjects will be consistent with previous studies</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>2. There will be a significant difference in response to the menthol test between subjects with knee OA and healthy controls</td>
<td>Accepted</td>
</tr>
<tr>
<td>3. There will be good associations between cold hyperalgesia as measured with ADI and cold hyperalgesia as measured with CPT.</td>
<td>Accepted</td>
</tr>
<tr>
<td>4. There will be significant differences in pain thresholds for cold, heat, pressure and punctate stimuli but no significant differences in sensory detection thresholds between subjects with knee OA and healthy matched controls.</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>5. There will be significant differences in physical function and quality of life between subjects with knee OA and healthy matched controls.</td>
<td>Accepted</td>
</tr>
</tbody>
</table>
7.6 Discussion

**Menthol test validity**

The primary aim of this study was to assess the ability of the menthol test to identify cold hyperalgesia in older individuals with and without chronic pain. Test validity was assessed by examining within-study associations and then comparing responses of healthy subjects from this study to those of Study 2 (Chapter 4) who were given the same gel.

Good significant correlations were found between the separate quality (MWS) and intensity elements of the menthol test for VAS cold, heat and unpleasantness (r= 4.18 to r= .541), suggesting internal construct validity. Intensity-quality associations for the same gel in Study 2 showed a similar correlation between unpleasantness VAS and the McGill descriptor score PRI (r= .563) (the ADI had not been available to use for Study 2). However, whereas in the current study VAS pain did not correlate with descriptor score, Study 2 showed a high correlation between VAS pain and PRI (r= .726). It is difficult to draw conclusions from this comparison as the descriptor scores used in the two studies were different, but it might be hypothesised that the older healthy subjects were more reluctant to describe the menthol sensation as painful. It is notable that only 11% of healthy older subjects reported >0 VAS in the current study compared with 53% of the younger healthy cohort in Study 2 (Chapter 4).

When comparisons were made between ADI scores and CPT temperature values, the current study demonstrated good concurrent criterion validity. Total ADI score and CPT temperature correlated at levels between r= .548 and r= .605 for all sites. MWS and CPT correlations were slightly stronger: from r= .560 - r= .647, following the trend seen in previous studies in this investigation for the best correlation to be between CPT and menthol descriptors. Overall, the menthol test in the current study appeared to show both good internal and external construct and criterion validity.

There were some consistencies in response to the gel between subjects in Study 2 (Chapter 4) and those in the current study. Although in both studies cold was experienced by at least 90% of all subjects, VAS heat, unpleasantness and pain were experienced by a considerably higher percentage of individuals from Study 2 than either the healthy or the OA subjects from the current study. Unpleasantness was reported by 11% of the healthy subjects from the current study, by 50% of the OA subjects but by 74% of Study 2 subjects. Level of intensity for each sensation was more
comparable, with healthy subjects from the current study rating all sensations only 5% lower than healthy subjects in Study 2, but OA subjects reporting 30-40% higher intensity, as might be anticipated. Similar words were used to describe sensation quality by all groups, although more Study 2 subjects selected more intense words such as burning or icy/freeing than either group from the current study.

The differences in healthy subjects’ responses may reflect slight differences in the test application. Although the presentation of words and general procedure was the same, VAS ratings were collected using e-VAS for the current study rather than the paper version for Study 2. However, very strong correlations have been found between touch-screen electronic VAS and paper VAS for reporting evoked stimuli (Jamison et al., 2002). There was also a difference in gel batch, although both were prepared from the same elements using the same procedure. Consistency of production is never the less an important factor to investigate in the future, since the margin of variability in gel preparation which will impact significantly on sensory response is unknown.

Alternatively the response difference between cohorts may reflect an age-related variability in pain response. Gagliese and Melzack (2003) reported significant age-related differences in choice of pain words from the MPQ, yet similarity in pain intensity between young (mean 43 years) and older (mean 70 years) subjects matched by pain diagnosis or location, pain duration and gender. The authors acknowledged the difficulty in concluding whether the differences reflect age-related differences in language use or actual pain experience but they propose that the latter is more likely. A recent fMRI study also supports the hypothesis that pain processes may become impaired with increasing age (Cole et al., 2010). An important follow-up study to the current investigation will be to clarify whether responses to the menthol test are indeed age-specific.

**Presence of cold hyperalgesia**

Response to menthol, tested at a forearm unaffected by osteoarthritis, clearly differentiated between healthy and osteoarthritic subjects, as proposed by Hypothesis 2. OA subjects recorded ADI scores on average 45% higher than their healthy counterparts. More than 50% of OA subjects found the sensation unpleasant and 20% found it painful, compared with 30% and 11% of control subjects. OA subjects also reported significantly higher intensities of cold and unpleasantness (Figure 7.3a-b) and were more likely to choose more intense or unpleasant descriptive words such as
icy/freeing, burning or prickling (Figure 7.6). The time-course of response also differed between OA and control groups. Whilst cold followed a similar pattern for both, the pattern for all other sensations was very different. For control subjects heat, unpleasantness and pain reached rapid peaks and then sharply declined to a minimal sensation well before 15 minutes (Figure 7.5a). In contrast, OA subjects reported a steady increase in heat, unpleasantness and pain, with the sensations only partially declining before 15-minute removal of the menthol stimulus (Figure 7.5b).

These intensity, quality and timing characteristics of response to menthol may indicate both the similarities and differences in cold sensation mechanisms between those with and without a painful pathology. The overwhelming response of all subjects to menthol was that of cold, experienced before any other sensation and by more than 90% of all subjects. This strongly supports previous studies in this investigation (Chapters 3 and 4) and indicates that menthol activates the cold receptor TRPM8 in both healthy individuals and those with chronic OA pain. Those with knee OA however reported almost double the cold intensity compared with control subjects. This may be the result of several potential mechanisms. It has been reported that individuals with OA may show signs of systemic inflammation which may cause a widespread increase in sensitivity: for example, levels of plasma serum C-reactive protein (CRP), which rise in response to inflammation, have been shown to be elevated in people with OA (Pearle et al., 2007). Inflammatory mediators sensitise Aδ and c-fibres, reducing thresholds to activation and increasing responses to evoked stimuli. Cold hyperalgesia has been shown in both human and animal models of arthritis (Allchorne et al., 2005; Moss et al., 2008). However, the role of TRPM8 during inflammation is unclear. On the one hand, there is evidence from animal models that TRPM8-null rats show reduced inflammatory-induced behaviour (Colburn et al., 2007). However, the balance of evidence suggests that inflammation both down-regulates expression of TRPM8 and reduces its activation thresholds. TRPM8 has been shown to be desensitised by bradykinin and prostaglandin PGE2 (Linte et al., 2007) and by substance P (Naono-Nakayama et al., 2009) and is also deactivated by the increased acidification that accompanies inflammation (Andersson et al., 2004). Human microscopy studies have shown significantly reduced TRPM8 nerve expression in those with increased cold sensitivity from pulpitis (Alvarado et al., 2007). It seems more likely that the increased intensity of cold sensation is the result of central sensitisation processes. Cold sensation is primarily transduced and transmitted via TRPM8 receptors expressed by Aδ thermo-receptors (McCoy et al., 2011). Central sensitisation in the dorsal horn of
the spinal cord may enhance the cold signal by a number of mechanisms. Central inhibition may be reduced, either as a result of deactivated GABAergic and glycinegic inhibitory interneurons or as a result of reduced descending inhibitory drive (or even enhanced descending facilitatory drive) (Basbaum et al., 2009). Increased excitability at central terminals as result of NMDA-mediated post-synaptic heterogeneous potentiation has been shown to increase transmission of non-noxious stimuli, such as cold in this case (Sandkuhler, 2009). Once the signal has been transmitted via the spinothalamic tract, additional changes in supra-spinal processing may also enhance the message. Although only a few imaging studies have been published that specifically explore functional brain changes in OA, the evidence suggests that enhanced supra-spinal processing of sensory information may occur in some people with OA. A recent, by Howard et al. (2011) reported increased resting cerebral blood flow in OA patients and Parks et al. (2011) found that spontaneous OA pain activated a range of prefrontal and limbic brain areas similar to that for post-herpetic neuralgia and chronic LBP.

In addition to increased cold intensity, larger numbers of subjects with OA reported significantly greater unpleasantness or pain alongside the cold during menthol application (unpleasantness: r = .444; pain: r = .438). Words such as burning or prickling were also selected more often by those with OA and the intense cold words icy/freeing were selected by 40% of OA subjects, compared with only 12.5% of controls. These findings indicate that, in addition to activation of non-noxious cold via Aδ-fibres, menthol evoked a nociceptive response in the OA group, suggesting that c-fibres and the spinothalamic pathway may have been activated. (Campero et al., 2009) has proposed that activation of high threshold c2-fibres may be responsible when burning sensations result from a cooling stimulus. These specialised c-fibres are normally silent but respond to cold (and intense heat) under certain circumstances. Temperature-cold has been used to show that an A-fibre block results in cool stimuli felt as unpleasantly icy, burning and stinging. During intense cold, these c-fibres, which are normally inhibited by cold-activated Aδ-fibre input, steadily activate whilst Aδ-fibres slow their activity (Campero et al., 2009). Unpleasant burning and other sensations are therefore delayed due to this gradual change in balance between c and Aδ output. Although the studies by Campero and colleagues did not examine cold receptor channels it is likely that TRPM8 was activated so that these result may be applied to a menthol model. C2-fibres required intense cold to start activating but central sensitisation would have the same effect by reducing thresholds to activation. Delayed activation of sensations of unpleasantness, icy, and burning were also paralleled in the current study.
Alternatively, it may be that noxious c-fibre stimulation is a reflection of activation of the receptor channel TRPA1, which is exclusively expressed on unmyelinated high-threshold c-fibres. TRPA1 has also been associated in human psychophysical studies with nociceptive qualities of unpleasant burning and stinging / pricking (Namer et al., 2005; Fajardo et al., 2008). Although it is a matter of great debate, Xiao et al. (2008) recently reported that TRPA1 is directly activated by menthol in humans, although not in the rat or mouse. TRPA1 has also been widely shown to be sensitised by a range of irritants including inflammatory mediators such as bradykinin, PGE2 and substance P (Bandell et al., 2004; Naono-Nakayama et al., 2009) and has been consequently termed “the gatekeeper” for chronic pain and inflammation (Bautista et al., 2012). Once activated, it has been proposed that TRPA1 may play a role in sensitising the high threshold TRPM8-expressing c-fibres that are usually kept silent by a potassium braking system (Serra et al., 2009; McCoy et al., 2011). It may therefore be that both TRPM8 and TRPA1 play a part in the unpleasant sensations of burning and pricking that were reported by subjects in this study. The significantly higher report of these qualities by OA subjects suggests that pathology-driven sensitisation may have had an influence.

**Additional signs of central sensitisation in OA subjects?**

Although not absolutely clear-cut, there were some additional signs that the OA subjects may have been experiencing some degree of central sensitisation. In addition to a cold hyperalgesia at the unaffected elbow, OA subjects exhibited significantly elevated cold pain thresholds (CPT) at the elbow and both knees. This supports the notion of a central effect, which may have been mediated at spinal or supra-spinal levels. The most elevated CPT temperature was seen at the OA knee, although this was not significantly different to the other sites. Pressure and punctate pain thresholds were also higher at both the OA knee and at the unaffected knee in OA subjects compared with controls. There was not a significant difference at the elbow, although PPT results from previous studies suggest that the control mean PPT may have been particularly low in the current study (Wylde et al., 2011) and a number of previous studies have found widespread increased sensitivity to mechanical stimuli in knee OA subjects (Imamura et al., 2008; Arendt-Nielsen et al., 2010). Although heat pain thresholds (HPT) were reduced (indicating sensitisation) at all sites for OA subjects, the difference was not statistically significant at the elbow and unaffected knee, although close to significance at the OA knee. Additionally, almost all detection
thresholds were reduced in the OA group compared with controls, with a statistically significant difference being seen just at the OA knee for cold detection. This finding largely supports other studies that have reported no increase in detection thresholds in people with OA, even if cold pain thresholds are elevated (Wylde et al., 2011).

Central sensitisation has been defined as the prolonged but reversible increase in excitability in central nociceptive pathways, “an amplification of neural signalling within the CNS that elicits pain hypersensitivity” (Woolf, 2011). Peripheral sensitisation describes the state whereby local inflammatory mediators create a sensitised environment so that innocuous stimuli activate nociceptors. Clinically, this is seen as a local hyperalgesia. Central sensitisation involves increased and extended excitability at the dorsal horn of the spinal cord, caused in part by homogeneous and heterogeneous potentiation. Once sensitised only a low level of stimulus is needed to reactivate a nociceptor; ‘silent’ nociceptors are activated and normally innocuous stimuli can trigger an entirely unrelated pathway. Clinically, central sensitisation is seen as widespread hyperalgesia, sensations that outlast the stimulus and allodynic or paradoxical responses that suggest system misinterpretation of the sensory input (Seifert & Maihofner, 2007; Gwilym et al., 2009).

Using these definitions, the current study found predominantly peripheral sensitisation in the OA cohort as a whole. Mechanical, punctate, heat and cold thresholds were elevated at the OA knee with little significant spread beyond the knee for all but cold hyperalgesia. Even though Lee et al. (2011a) reported a significant association between systemic CRP levels and widespread PPT, there is more evidence for the presence of inflammatory markers in serum CRP and knee synovium or cartilage (Pearle et al., 2007; Lee et al., 2011a). Local inflammatory mechanisms may therefore be proposed as the causative mechanism for peripheral sensitivity. Although the precise transduction mechanisms for mechanical hyperalgesia are still unclear, TRPA1 may be a strong candidate receptor. TRPA1 is required for mechanical and heat hyperalgesia (Bautista et al., 2012) and TRPA1 antagonists have been shown to attenuate mechanical hypersensitivity in a rat model of osteoarthritis (Eid et al., 2008). Other human models have shown that TRPA1 activation leads to both primary and secondary mechanical and pinprick hyperalgesia (Binder et al., 2011), as was shown in the current study. The fact that heat hyperalgesia is often reported in association with mechanical hyperalgesia during inflammation due to co-expressed TRPV1 and TRPA1 on c-fibres
may mean that this cohort of OA subjects did not have high levels of inflammation in their knee joint.

However, the clear presence of cold hyperalgesia at the elbow or forearm (increased CPT and ADI score) suggests involvement of some element of central pain augmentation in the OA group, which may originate from spinal or supra-spinal mechanisms. More frequent choice of words such as burning or prickling to describe a sensation that the majority of healthy subjects found cool or cold may suggest some centralised cross-talk or heterogeneous potentiation. The extended time-course of VAS heat, unpleasantness and pain sensations reported by OA subjects compared with controls may also suggest centrally-mediated potentiation of nociceptors or inhibitory mechanisms via c2-fibres. Additionally all of these findings may reflect supra-spinal changes in signal interpretation or in functional balance between descending inhibitory or facilitatory control.

It appears therefore that this cohort of OA participants did not exhibit highly augmented sensory hyperalgesia when compared with their control counterparts. Although there were clear signs of sensitisation localised to the OA knee and spreading to the contralateral knee, widespread hyperalgesia was limited to cold stimuli. The OA cohort in this study reported relatively mild mean pain and disability and this may explain the lack of greater sensitivity. On physical testing, there was a significant difference between OA and matched controls for each task and although OA subjects completed the tasks between 20-30% slower, the actual time difference was of little clinical significance (chair 1.2 seconds; walk 2.3 seconds; stairs 6 seconds).

However, it is also important to note that this study, in keeping with similar studies, used mean values to compare OA and control groups. A closer analysis of within-group variability shows that the OA group was not as homogenous as the control group (data collated into Table 7.10). So, although the OA group means in the current study were not always so different to those for the control subjects, this suggests that there were OA sub-groups, some of whom responded in a very similar manner to the controls whilst some exhibited a more strongly hyperalgesic response.
Table 7.10: Means, standard deviation, 95% confidence interval and variance for CPT values (°C) at the OA knee and elbow for OA and healthy subjects.

<table>
<thead>
<tr>
<th></th>
<th>OA subjects (n=40)</th>
<th>Healthy subjects (n=40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT OA knee</td>
<td>Mean 11.7</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>95% CI 8.7-14.6</td>
<td>1.9-5.3</td>
</tr>
<tr>
<td></td>
<td>SD 9.3</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>Variance 85.9</td>
<td>28.8</td>
</tr>
<tr>
<td>CPT Elbow</td>
<td>Mean 10.3</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>95% CI 7.8-12.8</td>
<td>1.6-4.9</td>
</tr>
<tr>
<td></td>
<td>SD 7.9</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>Variance 61.6</td>
<td>26.4</td>
</tr>
</tbody>
</table>

Study 8 sought therefore to investigate the differences between sub-groupings of OA subjects, using the data from 80 subjects with OA.
7.7 Summary

This study investigated the validity of the ADI as a test for cold hyperalgesia in both healthy subjects and those with painful knee OA. The menthol test showed good internal and external construct validity. Some differences were seen in intensity of response to the test by subjects in this study compared with younger healthy subjects in Study 2. This may be a factor of age-related differences in pain response (either in expression or in actual experience), which will need to be clarified by future studies.

This study also investigated whether matched OA subjects exhibited widespread hyperalgesia compared with their control counterparts. Widespread cold hyperalgesia was shown, with both forearm ADI and CPT clearly differentiating between OA and control subjects. The ADI showed significant group differences in total score, individual VAS intensity ratings and the percentage of subjects selecting specific words. Mean CPT temperature correlated well with ADI score. CPT was significantly different at both the OA affected knee and also at the unaffected knee and elbow. However, PPT and PcPT were significantly elevated only at the OA and contralateral knees but not at the elbow. Heat showed a trend towards significant increase at the OA knee but at no other site.

PPT, PcPT and HPT findings indicate more peripheral than central sensitisation effects in this particular cohort of OA subjects. However, widespread cold hyperalgesia found with both testing methods may indicate an underlying centrally-driven element to OA pain.

This study has shown that the menthol test is a valid instrument with which to identify individuals with cold hyperalgesia. Future studies need to investigate further the incidence and characteristics of centrally sensitised or augmented pain in a larger cohort with OA.
Chapter 8

Study Six

Comparison between knee osteoarthritis subjects with high and low ADI scores

8.1 Abstract

Background and Aims

It is now understood that knee osteoarthritis (OA) is not a homogeneous disorder that universally causes local nociceptive pain. Widespread hyperalgesia and neuropathic-type pain have been reported. Clinical studies have shown that standard analgesia and even joint replacement fails to reduce pain in 15-20% of patients. This study sought to investigate the association of widespread cold hyperalgesia, measured by forearm menthol ADI, with knee and elbow thermal and mechanical pain thresholds, incidence of neuropathic-type pain and severity of daily pain and dysfunction. ADI and CPT were compared to evaluate the efficacy of the ADI in identifying cold hyperalgesia.

Method

Eighty volunteers (36 male, 44 female, mean age 64 years) with painful knee OA were recruited. Following analgesic and NSAID wash-out, subjects underwent quantitative sensory testing with the menthol test applied at the elbow and mechanical, cold and heat hyperalgesia at both knees and one elbow in randomised order. Standard equipment, method and instructions were used. Self-reported pain, dysfunction and quality of life were assessed with WOMAC, PainDETECT, PQAS and SF36 questionnaires. The timed ALF assessed physical function.

Results

A good correlation was found between ADI and CPT temperature (r=.685, p<.001) and also between words chosen during menthol and CPT applications (r=.590, p<.001). 24% subjects were identified as cold hyperalgesic by both ADI and CPT. Following division into high and low ADI groups, a significant difference was seen for CPT and PPT at all sites (p<.001) and for HPT at the OA knee and elbow. PainDETECT was also significantly different (p<.001): 32% of the high ADI group scored as “positive neuropathic” compared with 5% of the low ADI group. The high ADI group also reported significantly higher levels of pain and dysfunction (WOMAC: pain p=.004; function p=.014).
Conclusion

There was a strong association between identification of cold hyperalgesia by menthol ADI or by thermal CPT. Subjects with ADI≥5 showed greater widespread mechanical and thermal hyperalgesia, higher levels of neuropathic-type pain and more severe pain and dysfunction. Widespread cold hyperalgesia was therefore associated with increased signs of central sensitisation and central pain augmentation.
8.2 Introduction & Background

The previous study (Chapter 7) compared the ADI and conventional quantitative sensory tests (QST) between individuals with knee OA and gender and age matched healthy controls. Although significant differences were seen for most measures between OA and controls, the study results suggested that the OA group was not homogeneous and that a sub-group with more elevated QST results exists. It has been proposed that this sub-group may have pain that is centrally-driven and so does not conform to the traditional nociceptive classification. Optimal treatment for this group is therefore likely to vary from standard interventions and so efficient and accurate identification would be paramount.

Conventionally osteoarthritis is seen as an archetypal “nociceptive” disorder, with pain driven by peripheral joint damage. The presence of inflammatory mediators such as nitric oxide and prostaglandin E2 (PGE2), the pro-inflammatory enzyme COX-2, and pro-inflammatory cytokines in the synovium and menisci of humans with knee OA is well documented (LeGrand et al., 2001; Brenner et al., 2004). Animal models of arthritis have shown that prostaglandins sensitise nociceptors provoking localised hyperalgesia and increased pain behaviours (Bove et al., 2006). A range of other pain-inducing structures in knee OA have also been shown to produce nociceptive output: periosteal tears, raised interosseous pressure, microfractures, and damage to ligaments, capsule or menisci (Felson, 2005). However, in human studies there is often a poor association between self-reported severity of pain and structural or biochemical markers of damage or inflammation. The poor correlation between radiographic disease severity and pain has been long acknowledged. Creamer et al. (1999) found that 40% of people with severe radiographic OA changes were pain-free. Brenner et al. (2004) found no significant associations between pain and any radiographic signs or levels of PGE2 or cytokines and Sengupta et al. (2006) found no association between MRI high-signal osteophytes and self-reported pain or pain location in people with knee OA.

Pain quality is an additional way to assess mechanisms and is even used, in the case of neuropathic pain, as an important diagnostic criterion (Huge et al., 2008; Baron et al., 2012). Studies comparing OA pain quality with other painful conditions report a tendency for those with OA to describe their pain in classical inflammatory terms such as aching and “throbbing” with episodes of “sharp” or “shooting” on movement (Hawker et al., 2008; Victor et al., 2008). Indeed, neuropathic pain questionnaires base
the validity of their ability to differentiate between nociceptive and neuropathic pain phenotypes on average OA pain quality data (Freynhagen et al., 2006; Victor et al., 2008). However, even in these validation studies, a proportion of OA subjects report classically neuropathic-type symptoms. Despite concluding that neuropathic qualities such as tingling or numb could be used to differentiate patients with carpal tunnel syndrome from those with OA, Victor et al. (2008) acknowledged that, for those with severe pain, this delineation may no longer apply. Recent studies have applied the PainDETECT neuropathic pain questionnaires to OA populations and found a significant proportion of subjects with scores in the positive neuropathic range (Hochman et al., 2011; Wylde et al., 2011). These findings suggest that a proportion of those with OA experience non-nociceptive symptoms so that it is no longer appropriate to consider the OA population as homogeneous.

It has been known for some time that a percentage of those with OA report poor pain efficacy with standard interventions. The percentage of those with ongoing pain after joint replacement surgery varies widely in the literature from 19-43% (Visser, 2006) and a number of studies have started to investigate the determinants of this failure. It has been proposed that there is a link between non-standard pain in OA and poor response to surgery. Valdes et al. (2012) found an inverse relationship between radiographic disease severity and ongoing pain severity after knee arthroplasty. Wylde et al. (2011) reported that 15% of individuals reported severe to extreme pain between 3-4 years post total knee replacement. This study found that 13% of those with any persistent pain scored as ‘positive neuropathic’ on the PainDETECT questionnaire. Using the same questionnaire with a community cohort, Hochman et al. (2011) found that 19% of people with knee OA scored in the positive neuropathic range and that ongoing pain and disability positively correlated with PainDETECT score.

In addition to self-reported non-nociceptive symptoms, additional signs of elevated non-localised pain have been reported, suggestive of centralised or centrally-augmented pain processing. A large number of studies have now reported signs of central sensitisation in knee OA, including secondary hyperalgesia (Bajaj et al., 2001; Arendt-Nielsen et al., 2010), widespread mechanical and heat hyperalgesia (Bradley et al., 2004; Imamura et al., 2008; Graven-Nielsen et al., 2012) temporal summation (Arendt-Nielsen et al., 2010). In contrast to biochemical and radiological studies, a good association has been found between widespread hyperalgesia and high pain levels and low quality of life (Imamura et al., 2008). Centrally-mediated dysfunction has also been
reported in people with OA, in particular a reduced endogenous pain inhibition (Kosek & Ordeberg, 2000; Graven-Nielsen et al., 2012). Imaging studies have reported grey matter changes in the thalamus and PAG brain biochemical abnormalities in those with painful OA (Gwilym et al., 2009; Gwilym et al., 2010).

Cold hyperalgesia has been proposed as a particularly important indicator of altered pain processing and associated with thalamic dysfunction (Seifert & Maihofner, 2007). Studies of musculoskeletal conditions such as whiplash have reported good associations between cold hyperalgesia and neuropathic pain as well as with mechanical hyperalgesia (Sterling & Pedler, 2009). Yet there is almost no data available regarding either localised or widespread cold hyperalgesia in people with painful OA, despite the large body of data demonstrating mechanical hyperalgesia.

This study therefore aimed to investigate whether widespread cold hyperalgesia, as measured by forearm menthol ADI score, was associated with measures of locally and centrally-driven pain in individuals with knee OA. Cold, mechanical heat and punctate thresholds were measured at local and distant test sites and data collected using both the PainDETECT and PQAS pain questionnaires. In addition, ADI and CPT were compared to evaluate the efficacy of the ADI in identifying cold hyperalgesia.

8.3 Hypotheses

1. There will be good associations between individuals exhibiting menthol-cold hyperalgesia (ADI ≥5) and thermal-cold hyperalgesia (CPT > 15°C).

2. Those with ADI ≥ 5 will report greater severity of pain, greater dysfunction and lower quality of life than those with ADI < 5.

3. Those with ADI ≥ 5 will perform functional tasks more slowly and with more pain than those with ADI < 5.

4. Those with ADI ≥ 5 will exhibit elevated widespread cold, heat, pressure and punctate pain thresholds, compared with those scoring ADI < 5.

5. Those with ADI ≥ 5 will report greater incidence of neuropathic-type pain symptoms compared with those scoring ADI < 5.
8.4 Methods

Subjects
Eighty subjects with painful knee osteoarthritis (OA) were voluntarily recruited from the Perth community via radio and newspaper advertisements. Volunteers were recruited and assessed for suitability by initial phone screen and then by a Rheumatologist, using the clinical criteria listed below (Table 8.1).

<table>
<thead>
<tr>
<th>Inclusion:</th>
<th>Aged 50 years +</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unilateral diagnosis of knee OA &gt; 6 months (American College of Rheumatology clinical classification system);</td>
</tr>
<tr>
<td></td>
<td>Pain in the index knee &gt; 3/10 on a pain visual analogue scale;</td>
</tr>
<tr>
<td></td>
<td>No additional clinically significant joint involvement;</td>
</tr>
<tr>
<td></td>
<td>In good health other than OA knee;</td>
</tr>
<tr>
<td></td>
<td>No arthroscopy or injections to index knee in last 6 months.</td>
</tr>
<tr>
<td></td>
<td>Able to read and understand English</td>
</tr>
</tbody>
</table>

| Exclusion: | History of systemic inflammatory conditions; |
|           | Neurological deficits (motor, cognitive or sensory); |
|           | Recent lower limb injury or surgery (< 6 months); |
|           | History of other chronic pain disorders (e.g. fibromyalgia); |
|           | Skin allergies; |
|           | Allergy to menthol. |

All participants provided written informed consent before participating in the study. Ethical approval was provided in a single submission for studies 5, 6 and 7 (Chapters 7-9) by Royal Perth Hospital Medical Research Ethics Committee (Approval EC 2009/100) and by Curtin University Human Research Ethics Committee (Approval number HR 26/2010).

OA participants using analgesics or non-steroidal anti-inflammatories (NSAIDs) underwent a washout period equal to five half lives of their analgesic or NSAID before testing. During the washout period participants were able to use acetaminophen for rescue analgesia but were asked to refrain from its use 12 hours before testing.
**Study Design and Procedures**

The study applied a cross-sectional design, analysing outcome measures assessed on a single test occasion. The 80 subjects, outcome measures and procedures were identical to those used in Study 7 (Chapter 9: 14-day placebo-controlled Arcoxia intervention) with this study analysing the baseline data from that study.

Following completion of informed consent procedures, subjects underwent a short interview to record any co-morbidities and usual medications. Subjects then completed the four self-report questionnaires for quality of life, pain and function and presence of neuropathic pain (see below for details). All subjects then completed physical and quantitative sensory testing (QST) outcome measures described below. Standardised method and instructions were used for each questionnaire and test. Order of physical and QST testing was randomised between modalities, although for punctate, heat and cold stimuli, detection threshold was always tested before pain threshold. The Aggregated Locomotion Function (ALF) tasks were always performed in the same order (chair, walk, stairs tasks).

**Outcome Measures**

QST and physical outcome measures were the same as for Study 5 (Chapter 3). Additional self-report pain and function measures were added in this cohort of subjects with OA pain.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QST</strong></td>
<td></td>
</tr>
<tr>
<td>Menthol- cold hyperalgesia</td>
<td>• Algotect Descriptor Index (ADI):</td>
</tr>
<tr>
<td></td>
<td>o VAS cold, heat, unpleasantness, pain;</td>
</tr>
<tr>
<td></td>
<td>o Pain quality descriptors (MWS)</td>
</tr>
<tr>
<td>Cold detection</td>
<td></td>
</tr>
<tr>
<td>Thermal -cold hyperalgesia</td>
<td>• Cold detection threshold (CDT) (°C)</td>
</tr>
<tr>
<td>Heat detection</td>
<td>• Cold pain threshold (CPT) (°C)</td>
</tr>
<tr>
<td>Heat hyperalgesia</td>
<td></td>
</tr>
<tr>
<td>Light touch detection</td>
<td>• Heat detection threshold (HDT) (°C)</td>
</tr>
<tr>
<td>Light touch allodynia</td>
<td>• Heat pain threshold (HPT) (°C)</td>
</tr>
<tr>
<td></td>
<td>• Punctate detection threshold (PcDT)</td>
</tr>
<tr>
<td></td>
<td>• Punctate pain threshold (PcPT)</td>
</tr>
</tbody>
</table>
Mechanical hyperalgesia | Pressure pain threshold (PPT) (kPa)

**Physical Function**
- Timed Aggregated Locomotion Function score (ALF)
- 2m return chair transfer
- 8m return walk
- Stairs

**Self-Report**
- Quality of Life
  - SF36-v2 Physical Health Score
  - Mental Health Score
- Pain and function
  - WOMAC OA Knee Index: pain, stiffness, function sub-scores
- Neuropathic pain
  - PainDETECT
- Pain quality
  - Pain Quality Assessment Scale (PQAS)

**Quantitative Sensory Tests:**
All quantitative sensory tests were applied at the same standardised sites as for Study 5. Menthol-cold hyperalgesia was tested at a 2x3cm site on the volar forearm 10cms above the wrist crease. Other QST tests were applied at the OA knee, the uninvolved knee and the ipsilateral elbow (Table 8.2). OA and "unaffected" knees were designated according to highest VAS pain rating.

<table>
<thead>
<tr>
<th>Punctate &amp; Pressure Thresholds</th>
<th>Thermal Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knee</strong></td>
<td>Medial knee joint line</td>
</tr>
<tr>
<td><strong>Elbow</strong></td>
<td>Common extensor origin</td>
</tr>
</tbody>
</table>

All QST tests were applied using the same equipment and standardised procedures as for Study 5 (Chapter 3). For all QST other than the menthol test, a practice was followed by 3 trials at each site, with the mean calculated for analysis.

For cold pain threshold testing (CPT), the same equipment and procedure were used as for Study 5 (Chapter 7). In addition, after each series of trials subjects were asked to rate the intensity and quality of pain experienced at CPT, by marking a paper 100mm VAS pain scale and selecting descriptors from the same list used with menthol testing.
**Menthol ADI Test:**

For the menthol test menthol gel B from the same formulation batch was used as for Study 5 (Chapter 7). It was applied to the forearm of all subjects, using exactly the same protocol as previously described. The same e-VAS and descriptor list measurement system as for Study 5 was also used, with ADI calculated as described previously in Chapter 7. All procedures were followed identically to study 5. In addition to calculating the overall Algotec Descriptor Index (ADI), the descriptor Mean Words Score (MWS) and values for maximum VAS intensity for each sensation, time to onset and time to peak for each VAS sensation were also recorded for analysis.

In order to divide subjects post-hoc into those with and without cold hyperalgesia, a cut-off value of 5 for the ADI score was used. The ADI was structured to identify and weight intensity ratings and abnormal sensory descriptors selected in response to a cold stimulus (Figure 8.1). Determination of normal and abnormal intensity ratings and descriptor selections was based on preliminary research using the McGill Descriptor Scale and VAS scales (Chapters 4, 5, 6 and 7).

The ADI score was structured to numerically combine both intensity values for sensations such as pain or heat and the selection of abnormal sensory descriptors, such as burning or stinging. The scoring system for each of these components was weighted so that a normal response to the stimulus (cold intensity VAS rating and selection of cold-related descriptors) would score ≤4 out of the total possible 9. To achieve a score of ≥5, an individual would need to select abnormal (high scoring) descriptors plus at least one abnormal VAS rating.

The validity of ADI 5 as a cut-off value was confirmed in Chapter 5, where a ROC analysis showed ADI predicted CPT group allocation with sensitivity .90, specificity .74 at a cut-off score of >4.9.
**Physical Function Testing**

The ALF timed physical function tests were also carried out in identical manner to Study 5 (Chapter 7), with each test completed three times and the mean time used for analysis. At the start and end of each test, subjects were asked to report the amount of pain from their knee, using a verbal VAS scale (0-10). Mean pain before, during and after each test was also calculated.

**Self-Report Questionnaires**

In addition to the SF36 v2 Health Survey and WOMAC knee index described in Chapter 7, this study used two measures of pain quality.

1. **PainDETECT** neuropathic pain diagnostic questionnaire is a validated self-report tool to identify neuropathic pain components in a range of conditions (Freynhagen et al., 2006). It has been applied in a number of studies to evaluate pain symptoms in OA (Gwilym et al., 2009; Hochman et al., 2011; Ohtori et al., 2012). The questionnaire uses a combination of VAS scale, body diagram and Likert-type questions asking about everyday frequency of symptoms such as ‘electric shocks’ or painful light touch. A total score is calculated, with subjects scoring less than 13/35

---

1 All permissions granted for use of SF36, WOMAC and PQAS questionnaires from copyright owners. PainDETECT is open access.
classified as 'negative neuropathic', 13-18 as 'unclear' and 19+ as 'positive neuropathic'.

2. Pain Quality Assessment Scale (PQAS) was also used to provide data regarding the type of spontaneous pain experienced by OA subjects (Victor et al., 2008). The questionnaire encompasses 17 numerical rating scales (NRS) numbered 0-10 each for a pain descriptor, plus additional NRS scales for unpleasantness and surface versus deep pain. The PQAS can then be calculated as a total or as four validated specific pain domains (Victor et al., 2008): global (intensity and unpleasantness), temporal (intermittent, variable, stable), paroxysmal (shooting, sharp, electric, hot, radiating), surface (itchy, cold, numb, sensitive, tingling) and deep (aching, heavy, dull, cramping, throbbing, tender). The subscales for paroxysmal, surface and deep will be calculated for this study as they may differentiate between neuropathic-type and nociceptive-type pain qualities.

Sample Size and Statistical Analysis
Based on previous studies (Hochman et al., 2011; Ohtori et al., 2012) it was predicted that PainDETECT and hyperalgesia measures such as the ADI would identify 20-30% OA subjects as experiencing more severe pain symptoms. Based on the estimated sample size of n=20 for the more severe group, it was calculated that the study would have 80% power to detect a 38kPa difference in PPT, a 5.4°C difference in CPT (Moss et al., 2007; Moss et al., 2008) and a 10.1mm difference in WOMAC (Moss et al., 2008) between groups.

Normality testing was carried out using a combination of Shapiro-Wilk tests and visual analysis of distribution. All data except CPT were normally distributed (Shapiro-Wilk p=.061 to .670) so that parametric statistics could be applied. CPT data required non-parametric statistical analysis when used as a single dataset. However, when divided according to CPT< or >15°C, each dataset was normally distributed, meaning that parametric statistics could be used.
Chapter 8

The following data analyses were carried out using the SPSS statistics package, version 19 with alpha set at p<.05.

<table>
<thead>
<tr>
<th>Hypothesis and Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Associations between menthol-cold hyperalgesia (ADI ≥5) and thermal-cold hyperalgesia (CPT &gt;15°C):</td>
<td></td>
</tr>
<tr>
<td>• % subjects recording ADI ≥5 and CPT &gt;15°C</td>
<td>Descriptive comparisons</td>
</tr>
<tr>
<td>• Correlation between ADI and mean global CPT temperature</td>
<td></td>
</tr>
<tr>
<td>• Correlation between menthol MWS and quality response at CPT (CPT-MWS)</td>
<td>Pearson’s or Spearman’s Correlation Coefficient</td>
</tr>
<tr>
<td>• Comparison between percentage of subjects selecting top ranked words during menthol and CPT stimuli</td>
<td></td>
</tr>
<tr>
<td>2. Differences between ADI &lt;5 and ADI ≥5 in self-reported pain, dysfunction and quality of life:</td>
<td></td>
</tr>
<tr>
<td>• VAS pain in last 7 days</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>• WOMAC-pain subscore</td>
<td></td>
</tr>
<tr>
<td>• WOMAC-function subscore</td>
<td></td>
</tr>
<tr>
<td>• SF36: physical (PCS) and mental health (MCS) values</td>
<td></td>
</tr>
<tr>
<td>3. ADI group differences in widespread cold, heat, pressure and punctuate pain thresholds between those with ADI &lt;5 and ≥5:</td>
<td></td>
</tr>
<tr>
<td>• CPT (°C)</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>• HPT (°C)</td>
<td></td>
</tr>
<tr>
<td>• PPT (kPa)</td>
<td></td>
</tr>
<tr>
<td>• PcPT (log value)</td>
<td></td>
</tr>
<tr>
<td>4. Group differences between ADI &lt;5 and ADI ≥5 in:</td>
<td></td>
</tr>
<tr>
<td>• PainDETECT score</td>
<td>Independent t-tests</td>
</tr>
<tr>
<td>• % subjects in each PainDETECT category</td>
<td></td>
</tr>
<tr>
<td>• PQAS paroxysmal, surface and deep sub-scores</td>
<td></td>
</tr>
<tr>
<td>5. Group differences in measures of physical function and quality of life:</td>
<td></td>
</tr>
<tr>
<td>• ALF: chair, walk, stairs tasks (secs)</td>
<td>Independent t-tests</td>
</tr>
</tbody>
</table>
8.5 Results

Baseline Characteristics of the OA cohort

- Gender, age and BMI
  The whole OA cohort comprised 45% male and 55% female subjects with a mean age of 64 years (range 50-86 years). Mean BMI was in the ‘overweight’ category, with 38% subjects registering in the ‘obese’ category (Table 8.3).

- Self-reported pain and function
  The OA cohort reported on average moderate pain and functional disability. Mean VAS pain for the previous seven days varied between 3/10 and 8/10 (mean 4.6). 53% subjects rated their mean pain as moderate (4-6/10) with 28% rating it as >6/10. WOMAC pain sub-scale ratings for specific knee-related activities mirrored this, with mean pain the equivalent of 3.7/10, with a maximum of 8.2/10 (scaled from raw WOMAC pain sub-scores /50). The majority of subjects also reported moderate to severe levels of stiffness (WOMAC stiffness: mean 8.2/20; max 18.9/20) and moderate levels of functional disability (WOMAC function: mean 53.4/250; max 123/170).

- Self-reported usual analgesia and co-morbidities
  The majority of subjects (65%) reported regular use of at least one analgesic medication. The majority used either slow release high dose acetaminophen ("Panadol Osteo") (40%) or NSAIDs (36%). Celebrex and stronger analgesics such as Tramadol were prescribed to only 10%, with opioids only reported by 2 subjects. Additional health issues that required medication were reported by 80% of subjects. The most prevalent conditions were diabetes (29%) and high blood pressure (29%), both controlled with medication. Intermittent low-grade low back pain (<3/10 average) was reported by 40% of subjects, with mild intermittent neck pain reported by 16% and migraines by 14% subjects. 14% of subjects reported taking medication for depression and 9% reported irritable bowel syndrome. All subjects were washed out of their usual analgesic and anti-inflammatory medications before testing but continued to take any other medications.

- Baseline data: ADI groups
  Once subjects were divided post-hoc according to ADI scores, although the low ADI group had a relatively equal gender mix, there was a significant difference between genders for the high ADI group, which comprised more than 70% females. There was no significant difference between ADI groups for age or for BMI.
Table 8.3: Comparisons between OA subjects with ADI < or ≥5 for gender, age and BMI.

<table>
<thead>
<tr>
<th></th>
<th>Total (n=80)</th>
<th>ADI &lt;5 (n=55)</th>
<th>ADI ≥5 (n=25)</th>
<th>t(74)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male: female)</td>
<td>36:44</td>
<td>29:26</td>
<td>7:18</td>
<td>$\chi^2=4.25$</td>
<td>.039*</td>
</tr>
<tr>
<td>Age: mean (range) yrs</td>
<td>65 (50-86)</td>
<td>64 (50-80)</td>
<td>67 (54-86)</td>
<td>-1.34</td>
<td>.183</td>
</tr>
<tr>
<td>BMI mean (SD)</td>
<td>28.6 (4.6)</td>
<td>28.3 (4.3)</td>
<td>29.2 (5.1)</td>
<td>- .92</td>
<td>.191</td>
</tr>
</tbody>
</table>

### Distribution of ADI Data

ADI analysis showed a distribution suggesting that the data was from a normally distributed sample group: Shapiro-Wilk p = .113.

![Histogram showing the normal distribution curve for ADI data from all subjects with knee osteoarthritis](image)

**Figure 8.2:** Histogram showing the normal distribution curve for ADI data from all subjects with knee osteoarthritis

### Hypothesis 1:

**There will be good associations between individuals exhibiting menthol-cold hyperalgesia (ADI ≥5) and thermal-cold hyperalgesia (CPT>15°C)**

- **There was a good correlation between ADI and CPT.**

There were good correlations across the whole OA cohort between ADI score (at the unaffected forearm) and CPT temperature at the same elbow ($r = .620$, p < .001) as well as at the OA knee ($r = .552$, p < .001) and for mean CPT temperature ($r = .685$, p < .001).

- **There was a good correlation between the quality of sensation experienced during menthol stimulus and at CPT**

There was a moderate correlation between descriptor score (MWS) for the menthol test and mean descriptor score for the experience at CPT (mean MWS across the three test sites): $r = .590$, p < .001.)
The words selected to describe the sensations experienced during the menthol stimulus and at thermal-CPT were also similar. As shown in Figure 8.3, cold or cool rather than icy/freezing tended to be the predominant temperature sensations during menthol stimulus. During thermal-cold CPT icy/freezing was more frequently reported. The more noxious words burning and penetrating were selected by similar percentages of subjects for both menthol and cold temperature stimuli.

**Figure 8.3:** Percentage of subjects selecting highest ranked words to describe sensations during the menthol stimulus and at CPT (mean word choice for the three sites).

- **Similar percentages of subjects scored in the high ADI and high CPT groups**

When subjects were divided dichotomously according to ADI, 25 subjects (31.3%) scored ≥5. Subjects were also grouped according to mean CPT temperature < or >15°C with 30 subjects (37.5%) classified in the high CPT group. If OA knee CPT was used as the grouping value, an additional 5 subjects recorded CPT>15°C. However, the same 19 subjects (24%) were classified in both high ADI and CPT groups whichever CPT value was used (Table 8.4).

**Table 8.4:** Membership of high and low ADI groups (< or ≥5) compared with membership of high and low mean CPT groups (< or >15°C average all sites)

<table>
<thead>
<tr>
<th></th>
<th>Low Mean CPT</th>
<th>High Mean CPT</th>
<th>χ²=23.00</th>
<th>p&lt;.001**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low ADI</td>
<td>44</td>
<td>11</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>High ADI</td>
<td>6</td>
<td>19</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

- **Individuals exhibiting CPT>15°C at all three test sites had the highest ADI scores and the highest PainDETECT scores (Figure 8.4).**

In order to explore widespread thermal cold hyperalgesia further, subjects were subdivided according to number of test sites with CPT>15°C: 42 subjects (52.5%) had
no test sites with CPT>15°C whilst 18 subjects (22.5%) exhibited CPT>15°C at all test sites. 10% of subjects exhibited CPT>15°C at only the OA knee. Independent t-tests showed that both forearm ADI and PainDETECT scores were significantly higher for those with widespread high CPT (CPT>15°C at all sites) compared with those whose CPT was <15°C at no sites (ADI: $t_{(58)}=-7.02$, $p<.001$; PainDETECT: $t_{(58)}=-5.60$, $p<.001$).

![Figure 8.4: ADI and PainDETECT scores for subjects grouped according to sites exhibiting CPT>15°C.](image)

- **Individuals exhibiting CPT>15°C at all test sites also reported the highest levels of pain and dysfunction.**

When those with CPT>15°C at all sites were compared with those exhibiting CPT>15°C at no sites, a significant difference was seen for mean VAS pain for the last 7 days ($t_{(58)}=-2.27$, $p=.027$) and also for WOMAC-pain ($t_{(58)}=-2.40$, $p=.020$).

**Hypothesis 2:**

**Those with ADI ≥5 will report greater severity of pain, greater dysfunction and lower quality of life than those with ADI <5.**

- **Those with ADI ≥5 reported significantly higher levels of pain and greater functional limitations, but no greater knee stiffness.**

There was a significant group difference in average pain (VAS) over the past 7 days: ADI<5 mean 4.3 ($±.19$) ADI≥5 mean 5.1 ($±.29$) $t_{(78)}=-2.20$, $p=.031$. Those with ADI ≥5 also scored significantly higher for pain during knee-specific tasks (e.g. going up stairs): WOMAC pain sub-score $t_{(78)}=-2.93$, $p=.004$ (Figure 8.3). WOMAC function score was also significantly different between those with ADI< or ≥5 ($t_{(78)}=-2.51$, $p=.014$).
However there was no significant group difference for WOMAC stiffness score ($t_{(78)} = -1.33, p = .188$).

![Figure 8.5: Mean scores for low and high ADI groups for WOMAC pain, stiffness and function sub-scores.](image)

- **Subjects with a high ADI reported significantly reduced quality of life due to physical problems but no reduction in quality of life due to mental health issues.**

There was a significant difference between ADI groups for SF36 physical sub-score ($t_{(78)} = 3.52, p = .001$). However there was no difference between groups for mental health sub-score ($t_{(78)} = 1.09, p = .281$).

**Hypothesis 3:**

**Those with high ADI scores will perform functional tasks more slowly and with more pain than those with ADI <5.**

- **Subjects with a high ADI score performed significantly less well and a higher percentage reported some pain during the ALF tasks. There was no significant group difference in pain intensity.**

There was a significant difference in ADI groups for time taken to complete each of the ALF tasks: chair transfer $t_{(78)} = -2.59, p = .011$; 8-metre walk $t_{(78)} = -3.48, p = .001$; stairs $t_{(78)} = -2.43, p = .017$ (Figure 8.6).

There also appeared to be a significant difference in pain report between groups both during and after almost all tasks (Figure 8.6). However there was also a significant group difference in pain intensity at the start ($p = .026$). Taking ADI group as the between-subjects factor, repeated measures ANOVA showed that there was no significant overall group difference in pain change between starting point and completion of the stairs ($p = .659$).
80% of subjects in the high ADI group reported pain >0/10 at the start with little fluctuation during the tasks and rest periods. For the low ADI group, there was greater fluctuation in pain reports. Although fewer subjects reported pain at the start, there was a progressive increase in those reporting pain during tasks. Whereas those in the high ADI groups felt pain regardless of task or rest, more of those in the low ADI group found that the pain brought on by a task dissipated immediately after the task was completed (Table 8.5). The stairs task elicited pain in almost all subjects.

### Table 8.5: Percentage of subjects in each ADI group reporting pain >0/10 VAS during and after each ALF task.

<table>
<thead>
<tr>
<th></th>
<th>ADI&lt;5 (n=55)</th>
<th>ADI≥5 (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>50.9</td>
<td>80.0</td>
</tr>
<tr>
<td>During chair task</td>
<td>70.9</td>
<td>84.0</td>
</tr>
<tr>
<td>After chair task</td>
<td>61.8</td>
<td>84.0</td>
</tr>
<tr>
<td>During walk task</td>
<td>74.4</td>
<td>88.0</td>
</tr>
<tr>
<td>After walk task</td>
<td>65.5</td>
<td>88.0</td>
</tr>
<tr>
<td>During stairs task</td>
<td>89.1</td>
<td>96.0</td>
</tr>
<tr>
<td>After stairs task</td>
<td>72.7</td>
<td>88.0</td>
</tr>
</tbody>
</table>

**Hypothesis 4**

*Those with ADI ≥5 will exhibit elevated widespread cold pain, heat pain, pressure pain and punctate pain thresholds, compared with those scoring ADI<5.*

- Subjects with ADI ≥5 showed significantly higher CPT at each site.

There was a significant difference in CPT values at both local and distant test sites between those with ADI <5 and those with ADI ≥5 (OA knee: t(70) = -3.83, p<.001; unaffected knee: t(70) = -6.54, p<.001; elbow: t(70) = -5.79, p<.001). For those with low
ADI, there was a significant difference in test sites ($F_{(2,48)} = 6.44$, $p = .002$) with CPT at the OA knee significantly higher. There was no significant difference in CPT between sites for those with ADI≥5 ($F_{(2,48)} = .375$, $p = .689$).

**Figure 8.7:** Mean cold pain threshold for low (<5) and high (≥5) ADI groups at each test site.

- **Those with high ADI selected more noxious words to describe the sensation at CPT**
  For each site subjects were asked to select words to describe the sensation at CPT. When subjects were divided according to forearm ADI, those in the menthol-cold hyperalgesic group reported higher levels of burning and stinging at CPT than those in the low ADI group. For each ADI group, the response at the OA knee was more intense than at other sites (Figure 8.8a-b).

**Figure 8.8a-b:** Most frequently chosen words at CPT for each test site: a) Non cold hyperalgesic subjects (forearm ADI <5); b) Cold hyperalgesic subjects (forearm ADI ≥5).

- **Those with high ADI also showed significantly lower (more sensitised) PPT values at each test site**
  There was a significant difference in PPT values at each test site between those with ADI <5 and those with ADI ≥5 (OA knee: $t_{(78)} = 4.64$, $p < .001$; unaffected knee: $t_{(78)} = 3.76$, $p < .001$; elbow: $t_{(78)} = 4.42$, $p < .001$).
Punctate pain threshold was only significantly different at the OA knee. There was a significant difference between ADI groups in PcPT at the OA knee ($t_{(78)} = -2.37, p = .020$) but not at the unaffected knee ($t_{(78)} = -1.64, p = .104$) or the elbow ($t_{(78)} = -.738, p = .463$).

HPT was significantly sensitised at the OA knee and elbow, although not at the unaffected knee for those with high ADI score. There was a significant difference in HPT at the OA knee ($t_{(78)} = 2.50, p = .015$) and at the elbow ($t_{(78)} = 3.68, p = .001$). Although there was a difference at the unaffected knee, it was not sufficient to reach statistical significance ($t_{(78)} = 1.76, p = .083$).

Hypothesis 5:
Those with ADI ≥5 will report greater incidence of neuropathic-type pain symptoms compared with those scoring ADI <5.

Those with ADI≥5 reported significantly greater neuropathic-type pain symptoms on PainDETECT.
There was a significant group difference in PainDETECT score ($t_{(78)} = -4.68$, $p < .001$) (Figure 8.11). Those with ADI <5 scored a mean of 10.2 (±.64) whereas those with ADI ≥5 scored an average 15.7 (±.87). When subjects were grouped according to PainDETECT categories, there was a significant difference in group membership ($p < .001$), with almost 1/3 of those with ADI ≥5 scoring in the “positive neuropathic” category (19+). In contrast only 5.4% of those with low ADI scores scored as "positive neuropathic" (Table 8.6).

**Table 8.6:** Cross-tabulation between membership of PainDETECT categories and ADI categories.

<table>
<thead>
<tr>
<th>PainDETECT categories</th>
<th>Negative &lt;13</th>
<th>Unclear 13-18</th>
<th>Positive 19+</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADI &lt;5 (n=55)</td>
<td>37 (67.3%)</td>
<td>15 (27.3%)</td>
<td>3 (5.4%)</td>
</tr>
<tr>
<td>ADI ≥5 (n=25)</td>
<td>5 (20%)</td>
<td>12 (48%)</td>
<td>8 (32%)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>42</strong></td>
<td><strong>27</strong></td>
<td><strong>11</strong></td>
</tr>
</tbody>
</table>

**Figure 8.11:** Mean pain questionnaire scores for low and high ADI groups: PainDETECT and PQAS paroxysmal, surface and deep sub-scores.

- Those with ADI ≥5 reported significantly greater neuropathic-type symptoms but no greater inflammatory-type symptoms with the Pain Quality Assessment Scale (PQAS). (Figure 8.11) There was a significant difference between ADI groups in scores for the paroxysmal (‘sharp’, ‘electric’, ‘radiating’) and surface (tingling, ‘itchy’, numb) sub-scores of PQAS (paroxysmal: $t_{(78)} = -2.29$, $p = .025$; surface: $t_{(78)} = -2.64$, $p = .010$). There was no difference in scores for the deep sub-score which asks about presence of ‘throbbing’, stinging and ‘cramping’ sensations ($t_{(78)} = -.362$, $p = .718$).
### Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Accepted/Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. There will be good associations between individuals exhibiting menthol-cold hyperalgesia (ADI ≥5) and thermal-cold hyperalgesia (CPT &gt; 15°C)</td>
<td>Accepted</td>
</tr>
<tr>
<td>2. Those with ADI ≥5 will report greater severity of pain, greater dysfunction and lower quality of life than those with ADI &lt;5.</td>
<td>Accepted</td>
</tr>
<tr>
<td>3. Those with ADI ≥5 will perform functional tasks more slowly and with more pain than those with ADI &lt;5.</td>
<td>Accepted</td>
</tr>
<tr>
<td>4. Those with ADI ≥5 will exhibit elevated widespread cold pain, heat pain, pressure pain and punctate pain thresholds, compared with those scoring ADI &lt;5.</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>5. Those with ADI ≥5 will report greater incidence of neuropathic-type pain symptoms compared with those scoring ADI &lt;5.</td>
<td>Accepted</td>
</tr>
</tbody>
</table>
8.6 Discussion
This study aimed to investigate the association between widespread cold hyperalgesia, as measured by forearm menthol ADI score, and additional measures of locally and centrally-driven pain in individuals with knee OA.

Validity of ADI as a measure of widespread cold hyperalgesia
It was proposed that the menthol ADI score could be used instead of more conventional CPT to identify those with and without cold hyperalgesia. Associations between ADI and CPT were therefore initially explored. Good positive correlations were seen between ADI score and CPT temperature at each test site: $r = .552$ to $r = .620$. The quality of sensation experienced (MWS descriptor sub-score) during the menthol test and during CPT was also well correlated ($r = .590$). Further break-down showed that specific words chosen during each type of cold stimulus were also similar, with the only clear difference seen in choice of words for the cold temperature aspect: menthol elicited cool or cold sensations whereas CPT elicited more intense icy/freezing sensations. Noxious words were selected relatively equally for both stimuli. These findings add further support to the notion that menthol is assessing the same phenomenon as CPT. Comparison of words chosen to describe the quality of sensation showed that noxious cold elicited a slightly more intense response overall. However, this is not surprising given that CPT word choice described the optimal cold pain stimulus for each subject whereas the menthol stimulus was standardised, meaning that for some it was a sub-optimal stimulus. If the word choice of those for whom the menthol stimulus was optimal (ADI score ≥5) are considered separately, 64% described the sensation as icy/freezing, which is more comparable to CPT. Study 5 (Chapter 7) has discussed the potential peripheral and central mechanisms by which both menthol and cold stimuli may elicit both cold and noxious sensations in individuals with a chronic pain pathology.

The percentage of subjects classified as cold hyperalgesic using ADI ≥5 cut-off or mean CPT >15°C were also similar at 31% and 37% respectively with 24% subjects classified as cold hyperalgesic by both methods. As an additional assessment of comparability between ADI and CPT as indicators of widespread cold hyperalgesia, subjects were classified according to the number of sites with CPT >15°C. Fifty-three percent of subjects did not exhibit CPT>15°C at any site whilst 22.5% exhibited CPT>15°C at every site. There were very strong associations between high CPT at all sites and high ADI score. When those with no cold hyperalgesic sites were compared with those who were
cold hyperalgesic at all sites, a clear pattern emerged. The ‘all sites’ group showed significantly higher ADI and PainDETECT scores and significantly lower PPT and HPT (all p<.001) (Figure 8.2). Self-reported pain (VAS and WOMAC) was also significantly higher in the ‘all sites’ group. This association between extent of spread of secondary cold hyperalgesia and severity of pain supports previous studies, which have found a similarly strong correlation between extent of secondary mechanical hyperalgesia and pain (Imamura et al., 2008; Arendt-Nielsen et al., 2010). This analysis also demonstrates that forearm ADI is comparable to CPT as an indicator of widespread cold hyperalgesia.

**The characteristics of cold hyperalgesic individuals**

Study 5 (Chapter 7) indicated that there was clear variability within the OA group in terms of their levels of pain, disability and hyperalgesia. Previous studies have found that a proportion of individuals with OA exhibit signs of altered pain processing (Imamura et al., 2008; Gwilym et al., 2010; Graven-Nielsen et al., 2012) and it has been proposed that the presence of widespread cold hyperalgesia may be a key indicator of such altered processing.

OA subjects in the current study were therefore divided according to ADI score with the cut-off set at ≥5 for widespread cold hyperalgesia. Group comparisons were made for pain, function and measures of hyperalgesia. This revealed significant group differences. Those individuals with widespread cold hyperalgesia also reported higher levels of spontaneous pain, both for VAS over the last seven days (p=.031) and for WOMAC-pain during knee-specific tasks (Figure 8.3). Everyday functional problems were also 37% worse in those with cold hyperalgesia and quality of life significantly lower. Physical performance was similarly more difficult for the cold hyperalgesic subjects, who completed all tasks significantly more slowly and with each task causing pain for more individuals (Figure 8.4 and Table 8.5).

Cold hyperalgesia was also associated with additional signs of widespread hyperalgesia and altered pain quality. Those with high ADI scores exhibited significantly greater hyperalgesia to cold, mechanical and heat stimuli at both local and distant test sites (Figures 8.5, 8.7 and 8.8). This was true not only when compared with values for OA subjects with low ADI but also when compared with mean values for healthy older subjects in Study 5. Table 8.7 lists comparative mean values.
**Table 8.7:** Comparison of cold (CPT), heat (HPT) and pressure (PPT) pain thresholds between older healthy subjects from Study 5 and OA subjects from the current study.

<table>
<thead>
<tr>
<th></th>
<th>Healthy older subjects (n=40)</th>
<th>Low ADI group (n=55)</th>
<th>High ADI group (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPT (°C)</strong></td>
<td>OA knee 2.6</td>
<td>9.4</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td>Unaffected knee 2.7</td>
<td>6.2</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Elbow 2.6</td>
<td>6.9</td>
<td>16.7</td>
</tr>
<tr>
<td><strong>HPT (°C)</strong></td>
<td>OA knee 45.5</td>
<td>45.9</td>
<td>44.1</td>
</tr>
<tr>
<td></td>
<td>Unaffected knee 45.9</td>
<td>45.9</td>
<td>44.6</td>
</tr>
<tr>
<td></td>
<td>Elbow 46.4</td>
<td>47.4</td>
<td>44.1</td>
</tr>
<tr>
<td><strong>PPT (kPa)</strong></td>
<td>OA knee 367.4</td>
<td>315.2</td>
<td>189.1</td>
</tr>
<tr>
<td></td>
<td>Unaffected knee 354.3</td>
<td>364.3</td>
<td>248.9</td>
</tr>
<tr>
<td></td>
<td>Elbow 310.9</td>
<td>375.8</td>
<td>263.7</td>
</tr>
</tbody>
</table>

ADI-cold hyperalgesic subjects also reported significantly greater neuropathic-type qualities to their spontaneous pain. Those with ADI ≥5 scored significantly higher on PainDETECT with 32% classified as ‘positive neuropathic’, in contrast to the low ADI groups, nearly 70% of whom scored as ‘negative neuropathic’. For PQAS, there were significant group differences in paroxysmal and surface sub-scores, both of which involve neuropathic-type descriptors.

Previous OA studies support these findings. A growing number of studies have reported widespread mechanical hyperalgesia in subjects with OA and a recent review by Suokas et al. (2012) has concluded that this phenomenon is a well-established characteristic in OA. Bradley et al. (2004) found decreased PPT at the shoulder in patients with knee OA whilst Wylde et al. (2010) and (2011) found significantly decreased PPT at the OA knee and forearm. Most recently, Harden et al. (2013) has added further support by reporting decreased PPT at the elbow in those with knee OA, compared with matched controls. Reduced pain, function and quality of life has been found to be associated with widespread mechanical hyperalgesia in several of these studies. Imamura et al. (2008) reported that mechanical hyperalgesia spreading to the sacrum in people with knee OA was associated with increased pain levels and reduced quality of life. Arendt-Nielsen et al. (2010) found an association between extent of widespread mechanical hyperalgesia (forearm and tibialis anterior at the shin) and pain severity.

A number of studies have also reported presence of neuropathic-type quality to spontaneous pain in subjects with OA. Hochman et al. (2011) found that 19% of those
with knee OA from a community cohort scored as “positive neuropathic” using PainDETECT. This study also found a strong correlation between self-reported pain and PainDETECT score, even once factors such as depression and pain catastrophising had been accounted for. Gwilym et al. (2009) grouped hip OA subjects according to high or low PainDETECT scores and found an association between the high score group (mean 17.7) and increased punctate pain threshold. However no additional PainDETECT information is reported. Ohtori et al. (2012) also reported a positive correlation between PainDETECT score and both pain VAS and WOMAC- pain scores. However, this study reported only 5.4% of subjects with “positive neuropathic” scores.

Previous studies which report the presence of widespread cold hyperalgesia in OA are more limited and the findings are unclear. Previous studies from our research group have found CPT significantly sensitised both locally and at distant unaffected heel and elbow sites in subjects with knee OA when compared with age-matched healthy controls (Moss et al., 2008)\(^1\). A similar study in subjects with hip OA found significantly sensitised CPT at the hip and also at the contralateral shoulder (Wright et al., 2010)\(^2\). In both of these studies, CPT difference was greatest at the affected joint. Other studies only partially support these findings. Harden et al. (2013) found that individuals with knee OA exhibited significantly higher CPT temperatures at the affected knee only when compared with controls. However, closer examination of data shows that whilst CPT values and standard deviations for the OA group were similar to the current study, those for controls seemed particularly high. Wylde et al. (2010) reported no differences in CPT between knee OA and control subjects at any sites, although OA subjects did exhibit reduced cold detection threshold, as with the current study. Specific CPT data are not available but Wylde reports use of the MSA (Somedic) thermode, which has a lower limit of 10°C which may have caused a floor effect. There are however a number of studies which have assessed cold hyperalgesia in different pain conditions and found significant sensitisation often associated with increased spontaneous pain or poor response to treatment. This is expanded upon below.

**Mechanisms: peripheral or central sensitisation?**

Harden et al. (2013) have recently proposed that publication of more QST data is needed to assist in understanding the heterogeneous mechanisms which underlie OA in particular whether peripheral or central sensitisation is involved. In the current study, those with ADI-cold hyperalgesia exhibited the highest levels of pain, disability and

\(^2\)Ref Appendix 5
quality of life. Those with high ADI scores exhibited an increased and noxious response to menthol at their unaffected forearm, which must involve central processing changes. This group also exhibited mechanical and thermal hyperalgesia that spread beyond the affected knee joint to the non-pathological upper limb site. So, although cold hyperalgesic subjects showed greater sensitivity to punctate pain at the OA knee only, the balance of evidence is towards central mechanisms as the predominant drivers of pain in this group.

Central sensitisation is an activity-dependent increase in excitation of second order neurons in the dorsal horn of the spinal cord. Multiple extra and intra-cellular processes are involved however there are several key players of relevance to the current study. Repetitive depolarisation at central interneurons and spinal projection neurons causes increasing levels of pro-nociceptive molecules such as substance P and glutamate. This is thought to gradually dislodge the magnesium ion block that normally inhibits activation of post-synaptic NMDA receptors (Latremoliere & Woolf, 2009). The resulting influx of calcium ions leads to phosphorylation of the NMDA receptor followed by multiple subsequent down-stream events that produce sustained levels of excitation in the dorsal horn. Activation of NMDA is therefore a key trigger for long term potentiation and amplification of nociceptive input. Receptors such as TRPA1 are more easily activated by heat, cold and mechanical input and previously ‘silent’ nociceptors are activated. Concurrent changes in inhibitory mechanisms in the dorsal horn add to the enhanced effect. Nociceptive input is normally controlled by tonic input from GABA-ergic and glycinergic inhibitory interneurons. However, during central sensitisation these interneurons are deactivated by loss of GABA and glycine in response to increases in expression of pro-nociceptive chemicals (Basbaum et al., 2009). In addition, spinal inhibition of incoming nociceptive messages may be reduced by supra-spinal changes in the balance between facilitatory or inhibitory descending influence. During central sensitisation (or possibly as a pre-existing phenomenon for some) descending inhibitory influence is reduced and additional facilitation of nociceptive input is triggered (Latremoliere & Woolf, 2009). As a result of combined increased spinal excitation and decreased inhibition, nociceptive signals are augmented and intensified. This is experienced as hyperalgesia to evoked stimuli or as increased intensity of spontaneous pain (Sandkuhler, 2009). Increased central excitation also results in increases in receptive fields plus homosynaptic and heterosynaptic potentiation which causes abnormal ‘cross-talk’ between neurons so that normally non-nociceptive afferent information is coded as painful. For example, secondary
hyperalgesia describes the increased sensitivity to a stimulus of non-noxious intensity at an undamaged site distant from but within the same region as the painful tissue. In the current study, this was seen as mechanical and heat hyperalgesia at the unaffected knee. The fact that hyperalgesia was also observed at the upper limb site suggests that central sensitisation effects were also influenced by centrally-mediated changes to descending inhibition. Although not specifically tested with QST, the current study also showed evidence of functional pain summation and longer-lasting pain, both of which are described as features of a centrally sensitised state. The ALF tasks caused a steady increase in pain which, for the cold hyperalgesic group did not abate immediately after task completion (Figure 8.4). Given that pain that outlasted the stimulus was seen in those in the high ADI group, this may indicate that widespread cold hyperalgesia identifies those with altered pain processing that directly impacts on everyday activities, not just on experimental evoked pain states.

Central sensitisation that is maintained by ongoing nociceptive input has been shown to be reversible. Harden et al. (2013) have suggested that OA may be a particularly robust model for activity-dependent central sensitisation since the disorder involves ongoing barrage of pain from structures within and around a large and clearly damaged joint and possibly also from damaged neural structures. The current study found that those with widespread hyperalgesia showed the greatest intensity levels at the affected knee joint. This included greater intensity of burning and stinging pain quality (Figure 8.6b) and suggests that peripheral nociceptive input was at least partly responsible for the central sensitisation. A number of studies have demonstrated that even apparently long-established signs in those with knee OA are reversible following removal of the damaged joint. Kosek and Ordeberg (2000) demonstrated in subjects with hip OA that the mechanical hyperalgesia shown in both the operated and contralateral hip preoperatively normalised post arthroplasty. DNIC mechanisms were also normalised. Graven-Nielsen et al. (2012) showed that self-reported pain plus PPT at the knee, forearm and tibialis anterior muscle and conditioned pain modulation (CPM), the observable sign of descending inhibition, returned to normal levels by 9-18 weeks following knee joint replacement surgery. This has even been shown for changes in grey matter volume which returned to normal levels nine months following hip joint replacement (Rodriguez-Raecke et al., 2009; Gwilym et al., 2010)
**Mechanisms: central pain augmentation?**

However there is evidence that not everyone with OA has pain which resolves once nociceptive input is removed, suggesting that more established supra-spinal mechanisms may perpetuate the pain signal for some. Kehlet H et al. (2006) reported that 5–10% of those undergoing surgery will develop moderate to severe persistent pain. Wyld et al. (2011) found that 3-4 years post surgery, 15% of patients with total knee (TKR) and 6% of those with total hip replacements reported regular severe to extreme pain. Of these, 13% of TKR and 5% of THR patients reported sufficient neuropathic-type symptoms to be classified as ‘positive neuropathic’ using PainDETECT.

It has been proposed that report of neuropathic pain-type symptoms in pain conditions that have no identifiable neuropathy is a sign of more established central pain processing abnormalities. Positive score on PainDETECT has been reported in those experiencing persistent pain following thoracic surgery. Steegers et al. (2008) found that there was a strong positive correlation between PainDETECT score and intensity of chronic pain, with 23% of those with persistent pain scoring as ‘positive neuropathic’ and a further 30% scoring as ‘unclear’. In fibromyalgia, PainDETECT score has been associated with higher self-reported pain and increased tender point count (Amris et al., 2010). Ohtori et al. (2012) have reported that, although less than six percent of subjects with OA knee scored as ‘positive neuropathic’, there was a significant positive correlation between PainDETECT and self-reported pain but a negative correlation between PainDETECT and radiographic or biochemical signs of OA. In other words, neuropathic-type pain was associated with fewer signs of nociceptive input. In the current study, PainDETECT score was significantly higher in those with ADI-cold hyperalgesia compared with the low ADI group. Thirty-two percent of those with ADI≥5 were classified as ‘positive neuropathic’. Pain quality during both forearm menthol testing and at CPT was reminiscent of neuropathic qualities, with cold hyperalgesic subjects more likely to select burning and stinging.

Neuropathic pain quality may of course reflect nerve damage. Sterling and Pedler (2009) found that persistent whiplash pain was associated with neuropathic score on the self-reported Leeds Assessment of Neuropathic Symptoms and Signs Scale (sLANSS) and concluded that undiagnosed small fibre nerve damage was the most likely explanation. Nerve damage seems equally possible for OA where inflammation may sensitise nerves, or where joint damage may more permanently damage neural
structures. Neuropathic damage is described as producing both positive and negative effects. In the current study, whilst those with positive PainDETECT scores reported spontaneous tingling and burning sensations, there were no signs of globally reduced function. No group differences were seen in punctate detection or heat detection thresholds and although there was a significant ADI group difference in CDT this was driven by low detection thresholds in the ADI<5 group. Mean CDTs for the cold hyperalgesic group were very similar to those for healthy older subjects in Study 5 (Chapter 7).

Spontaneous neuropathic-type symptoms may instead reflect central augmentation of pain. At a spinal level unpleasant dysaesthetic sensations, hyperalgesia and alldynia may be relatively hard-wired and involve transcription changes or even phenotypic switching (Costigan et al., 2009). Equally, clearly defined neuropathic-pain type symptoms may be an indication that more hard-wired changes have taken place at a cortical or sub-cortical level. fMRI studies (Gwilym et al., 2009; Gwilym et al., 2010) have demonstrated volume and activity changes in thalamic grey matter and in the mid-brain periaqueductal grey (PAG) in people with OA hip pain associated with activation of the PAG. This clearly indicates that central brain changes are involved in the mediation of neuropathic-type symptoms in people with OA. Changes in neurochemical pathways and structural reorganisation in the pre-frontal cortex and thalamus have also been demonstrated in patients with fibromyalgia (Gracely & Ambrose, 2011) and chronic low back pain (Apkarian et al., 2011).

Widespread cold hyperalgesia has also been proposed as a key indicator of more established supra-spinally augmented pain processing (Stone et al., 2012). In the current study, CPT was strikingly consistent between sites. Those with high forearm ADI scores exhibited thermal-cold hyperalgesia (CPT>15°C) at all test sites and CPT temperatures were consistently between 16.7°C and 17.8°C. This is in contrast to CPT values of 9.4°C (OA knee), 6.2°C (unaffected knee) and 6.9°C (elbow) for the low ADI group. This consistency for both sensitised and non-sensitised subjects suggests strongly that a central mechanism is likely be controlling this response. A few other studies have provided evidence that points to cold pain sensitivity as an indicator of centrally-driven processes and poor response to interventions. In a range of studies investigating whiplash associated disorder (WAD), elevated cold pain threshold has been clearly shown to be a strong predictor of pain chronicity, severity and dysfunction (Sterling et al., 2006; Sterling et al., 2012) and associated with poor response to
treatment (Jull et al., 2007). In experimental studies, Tuveson et al. (2003) concluded that cold pain sensitivity that was of higher intensity and lasted longer than equivalent heat pain sensitivity following hypertonic saline injection in healthy subjects was a stronger sign of centrally-driven sensitisation. A number of studies have shown that cold hyperalgesia is strongly linked to chronicity and severity in fibromyalgia, the foremost "central sensitivity syndrome". Hurtig et al. (2001) found that CPT was positively associated with higher pain intensity, increased tender points and decreased sleep quality. Smith et al. (2008) found that female subjects with fibromyalgia showed sensitisation to repeated cold stimuli in contrast to habituation by control subjects, again suggesting a more entrenched phenomenon. An early QST study showed increased sensitivity to non-painful cold (often experienced as paradoxical heat) at all body sites, in contrast to heat hyperalgesia only at pain sites Kosek et al. (1996). None the less, the evidence for cold hyperalgesia as a key indicator of pain processing is still very limited and the exact mechanisms by which this might happen are largely unknown.

**Future analysis of mechanisms**

One approach to clarifying the mechanism driving the widespread hyperalgesia and neuropathic-type symptoms experienced by 30% of OA subjects in the current study would be to evaluate whether a known pharmaceutical intervention could influence these symptoms. A recent study (Chappell et al., 2009) has found that duloxetine, a centrally acting serotonin and norepinephrine reuptake inhibitor was significantly more successful than placebo in reducing pain and improving function in subjects with painful knee OA. Although the central action of duloxetine was effective, no specific measures of neuropathic pain or hyperalgesia were used in this study and so questions about central mechanisms remain unanswered. A converse approach would be to study the efficacy of a medication that will reduce peripherally-driven nociception but that is not primarily centrally-acting.

The final study in this investigation (Chapter 9) therefore sought to investigate the efficacy of the cyclo-oxygenase 2 inhibitor (COX-2) etoricoxib (Arcoxia) in influencing measures of centrally-mediated pain in individuals with knee OA. The effectiveness of etoricoxib on QST measures of local and widespread hyperalgesia and incidence of neuropathic-type symptoms was investigated.
8.7 Summary

This study aimed to investigate the association between widespread cold hyperalgesia, as measured by forearm menthol ADI score, with additional measures of local and central pain in individuals with knee OA.

Comparison between forearm ADI score and CPT temperature showed that the menthol test was equally effective as CPT in identifying cold hyperalgesia in this clinical cohort. Correlations between ADI and CPT were good and descriptions of sensation quality was similar. There was also a clear cross-over in membership of the thermal-cold hyperalgesic and the menthol-cold hyperalgesic groups. An ADI cut-off of 5 was therefore used to differentiate OA subjects into those with and without widespread cold hyperalgesia, resulting in a cold hyperalgesic group of 25 subjects (31.3%).

This cold hyperalgesic group exhibited significantly increased levels of self-reported pain, disability and quality of life compared with the non cold-hyperalgesic group. Physical function performance was also reduced in the hyperalgesic group, with slower times and higher percentage of subjects reporting pain during chair, walk and stairs tasks. Cold hyperalgesic subjects also exhibited mechanical and heat hyperalgesia at both the local OA site and at the distant elbow sites. Higher levels of neuropathic pain during everyday life were also reported by the high ADI group, with 32% classified as ‘positive neuropathic’ according to PainDETECT score in contrast to only 5% of those with low ADI. Results for PQAS matched this.

These findings may therefore reflect predominantly centrally-sensitised characteristics in those OA subjects with ADI-cold hyperalgesia. Widespread cold hyperalgesia and neuropathic-type sensory quality to both spontaneous and evoked pain may suggest an additional element of supra-spinal pain augmentation.

However, further studies are needed to further elucidate whether spinal or cortical mechanisms are the primary drivers of persistent pain in those with widespread cold hyperalgesia.
Chapter 9

Study Seven

The effect of etoricoxib on hyperalgesia, pain and function in subjects with knee OA

9.1 Abstract

Background and Aims
Cross-sectional studies have shown that knee osteoarthritis (OA) is a heterogeneous disorder with varying pain drivers, which may generate differential responses to interventions. This double-blind, randomised, placebo-controlled study sought to investigate response to a standard Cox-2 inhibitor (etoricoxib, 60mg daily) over 14 days. The menthol test was applied alongside additional quantitative sensory tests (QST) for thermal and mechanical hyperalgesia, PainDETECT and measures of pain and dysfunction. Comparisons were made between those with and without widespread cold hyperalgesia and the sensitivity of the menthol ADI test evaluated.

Method
Eighty volunteers (36 male, 44 female, mean age 64 years) with painful knee OA were randomly allocated to placebo or active groups. Subjects were assessed at baseline, day 4 and day 14 for mechanical, cold and heat hyperalgesia at both knees and the elbow. Standard equipment, method and instructions were used. The menthol test was applied at the forearm. Self-reported pain and dysfunction was assessed using WOMAC, PainDETECT and PQAS questionnaires. The timed Aggregated Locomotion Function (ALF) score assessed physical function.

Results
There were significant group x time interactions for all measures of pain, dysfunction and quality of life: Active group improvements of 30% for WOMAC pain and function, and 10-40% reductions in ALF speed and pain. Significant group x time interactions were also seen for stiffness and hyperalgesia at the OA knee (WOMAC stiffness p=.001; PPT p=.001). CPT at the OA knee significantly improved for the Active group (p=.037), and there were significant group x time interactions for widespread measures: ADI, PainDETECT and all PQAS sub-scores p<.001. When subjects were divided according to presence of cold hyperalgesia (ADI< or ≥5 at Day 0), those with low ADI showed clearer improvements in PPT, WOMAC and PQAS sub-scores by Day 14. Changes in PainDETECT and ADI were significant for both ADI groups (p<.001), and CPT non-
significant for both. Effect size coefficient for ADI approached that for PainDETECT but surpassed that for CPT ($r = .444$; $r = .581$; $r = .111$ respectively).

**Conclusion**
Etoricoxib influences both local and widespread pain measures, although is more clearly effective in non-cold hyperalgesic individuals. The menthol ADI is a sensitive alternative measure for widespread cold hyperalgesia.
9.2 Introduction & Background

The previous study in this investigation provided evidence that are sub-groupings of individuals exhibiting clusters of spontaneous and evoked pain responses associated with varying severity of dysfunction. Study 6 also demonstrated that widespread cold hyperalgesia, as defined by forearm ADI score for those with knee OA, can be used to identify a group that experiences more comprehensively elevated pain and dysfunction. It has been proposed that this more severe pain presentation is reflective of centrally-augmented pain, driven by CNS altered processing rather than by intensity of peripheral nociceptive input and that these more severely affected individuals may respond poorly to standard interventions such as NSAIDs or surgery. This study applied a placebo-controlled design to assess whether modulating inflammation with a Cox-2 inhibitor would influence measures of sensitisation and whether it would have a differential effect on those identified as having more severe pain and dysfunction associated with the presence of cold hyperalgesia.

Determining the relative contributions of peripheral or central mechanisms to chronic pain is challenging. Studies of amputees who experience pain in the amputated limb demonstrate that that ongoing peripheral nociceptive input is not necessary for pain to be felt in a particular anatomical location. For OA, there is an ambivalent link between radiological severity and pain severity, indeed it has been reported that up to 10% of those with severe self-reported pain and disability show no radiological signs of OA (Dieppe et al., 1997). Biomarker studies are equally unclear in showing an association between signs of an inflammatory or immune response and subjective report of pain in OA (Brenner et al., 2004; Lee et al., 2011). Knee joint replacement is the ultimate test of whether nociceptive input is driving pain. It has been reported that up to 30% of patients report ongoing pain in their knee several years after surgery, and several studies have found an inverse relationship between radiological severity pre-op and ongoing pain severity post-op. This appears to indicate that for some individuals their pain is driven by alterations in central processing rather than peripheral nociceptive input.

However, joint replacement result aside, additional measures may reflect peripheral versus central mechanisms. Local hyperalgesia, particularly to mechanical, punctate and heat stimuli, may identify reversible peripheral sensitisation. Local inflammatory-induced sensitisation may also be reflected in self-report of knee stiffness alongside inflammatory-type pain qualities such as throbbing and aching (Victor et al., 2008). In
Chapter 9  Effect of etoricoxib in subjects with knee OA

contrast, the presence of widespread hyperalgesia, as well as temporal summation and increased receptive fields (Arendt-Nielsen et al, 2010) have been proposed as measures of spinally-driven central sensitisation. More established changes in CNS pain processing at a cortical level have been assessed using imaging or psychophysical tests of descending pain modulation efficiency (Gwilym et al., 2009; van Wijk & Veldhuijzen, 2010).

Widespread cold hyperalgesia has also been proposed as a particularly robust sign of more established altered central pain processing. Persistent cold hyperalgesia has been associated with increased severity in fibromyalgia (Hurtig et al., 2001) and poor treatment outcome in whiplash (Jull et al., 2007), possibly associated with thalamic dysfunction (Kim et al., 2007; Lindstedt et al., 2011) or changes in central control of pain inhibition (Kosek & Ordeberg, 2000). It may therefore be that widespread cold hyperalgesia is an indication of processes that are relatively independent of nociceptive input. Cold hyperalgesia is associated with altered pain qualities, such as a paradoxical burning sensation, which may also signal neuroanatomical changes in the pain interpretation system, either at spinal level or at higher centres (Seifert & Maihofner, 2007). Finally, neuropathic pain questionnaires have been used to identify people with musculo-skeletal conditions who report non-nociceptive-type qualities to their pain (Sterling & Pedler, 2009; Wylde et al., 2011; Hochman et al., 2012) and it has been suggested that, rather than identifying undiagnosed neuropathy, a positive PainDETECT score may reflect centrally-augmented pain processing (Amris et al., 2010; Gwilym et al., 2010).

However, it is by no means clear that this framework of local versus widespread hyperalgesia, response to cold and neuropathic pain questionnaires does in fact differentiate between peripheral sensitisation, central sensitisation and more established centrally-augmented pain. An additional approach is to use a pharmaceutical intervention with a known peripheral or central action and evaluate change in these suggested outcome measures. A widely-referenced study of duloxetine has shown that a centrally-acting serotonin-norepinephrine reuptake inhibitor has a positive impact on knee OA pain and function (Chappell et al., 2009). However this study focused on general efficacy rather than mechanisms and so only used conventional self-report outcome measures with no sub-division of subjects according to pain presentation. This study instead proposed to assess the effects of an anti-inflammatory intervention on local and widespread psychophysical measures and self-
reported neuropathic pain. An anti-inflammatory would primarily influence the local environment, which may also involve central spinal sensitisation but would be unlikely to influence signs of more hard-wired CNS changes in pain processing.

The Cox-2 inhibitory anti-inflammatory etoricoxib was used in this study. Cox-2 inhibitors target more specifically the cyclo-oxygenase 2 enzyme which catalyses arachidonic acid in the first stage of its conversion to pro-inflammatory prostaglandins. Prostaglandin 2 (PGE2) plays an important role in inflammatory sensitisation, for example increasing the sensitivity of sodium channels leading to increased discharges from sensory and nociceptive neurons. COX-2 inhibitors such as etoricoxib have been shown in both animal and human models to be effective in reducing catabolism of arachidonic acid by inhibiting the activity of the COX-2 enzyme (Bingham et al., 2006). Cytokine pathways create a complex positive feedback loop so that increased production of PGE2 causes increased production of Cox-2. Both Cox-1 and Cox-2 have been shown in animal models to be strongly upregulated during inflammation (Bingham et al., 2005). It has therefore been proposed that Cox-1 plays a greater role in acute inflammation whereas Cox-2 is more involved in controlling more chronic inflammatory states (Bingham et al., 2006). Several studies have reported the efficacy of etoricoxib in reducing inflammation (Renner et al., 2010) and improving self-reported pain and function (Lin et al., 2010). However no previous etoricoxib studies have applied QST or neuropathic pain measures and none to our knowledge have analysed response of different pain groups.

Etoricoxib (Arcoxia) is licensed by the Australian Therapeutic Goods Administration for OA at 30mg and 60mg daily doses. A number of studies have shown that the 60mg daily dose has improved efficacy whilst limiting cardiovascular risk in knee and hip OA when taken over longer periods of time (Bingham et al., 2008). Doses of 120mg are more appropriate for short-term use, for example in post-operative pain (Boonriong et al., 2010). This study used a 60mg daily dose of etoricoxib for those in the Active group. Those in the Placebo group were provided with acetaminophen as rescue medication but were requested to refrain from taking any 12 hours before their testing sessions. Given that the study aimed to evaluate whether a Cox-2 reduced inflammatory-associated nociception it was important to ensure that Placebo results reflected the full extent of nociceptive input experienced by the group. All subjects were washed out of any usual analgesics or NSAIDs before the first test session, again in order to limit confounding variables. Subjects and assessors remained blind to group throughout the
study since a number of studies have found that expectation is a strong mediator of effect with many types of medical intervention (Tracey, 2010).

This study therefore applied a double-blind placebo-controlled design to assess whether modulating inflammation with a Cox-2 inhibitor (etoricoxib) over two weeks would be effective in reversing indicators of peripheral and central sensitisation and centrally-augmented pain as well as local hyperalgesia in people with knee OA. In addition, the study explored whether efficacy was reduced in those individuals who exhibit widespread cold hyperalgesia.

9.3 Hypotheses

1. When compared with subjects receiving a placebo formulation, subjects with knee OA who receive a daily dose of etoricoxib 60mg for 14 days will demonstrate:
   a. A significant improvement in pain, dysfunction and quality of life when compared with subjects receiving placebo formulation.
   b. A significant reduction in measures of local hyperalgesia at the OA knee when compared with subjects receiving placebo formulation.
   c. A significant reduction in widespread hyperalgesia when compared with subjects receiving placebo formulation.
   d. Significant improvements in self-report score for neuropathic pain when compared with subjects receiving placebo formulation.

2. The measures of cold pain threshold and ADI will show similar responsiveness to the intervention of etoricoxib 60mg for 14 days.

3. Subjects with cold hyperalgesia (ADI ≥5) will respond less well to etoricoxib than those with a less hyperalgesic response (ADI<5).
9.4 Methods

Subjects

Eighty subjects with painful knee osteoarthritis (OA) were voluntarily recruited from the Perth community via radio and newspaper advertisements. Volunteers were assessed for initial suitability by phone screen and then assessed for medical suitability by a Rheumatologist, using the criteria listed below (Table 9.1).

<table>
<thead>
<tr>
<th>Inclusion:</th>
<th>Aged 50 years +</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unilateral diagnosis of knee OA &gt; 6 months (American College of Rheumatology clinical classification system);</td>
</tr>
<tr>
<td></td>
<td>Pain in the index knee &gt;3/10 on a pain visual analogue scale;</td>
</tr>
<tr>
<td></td>
<td>No additional clinically significant joint involvement;</td>
</tr>
<tr>
<td></td>
<td>In good health other than OA knee;</td>
</tr>
<tr>
<td></td>
<td>No arthroscopy or injections to index knee in last 6 months.</td>
</tr>
<tr>
<td></td>
<td>Able to read and understand English</td>
</tr>
</tbody>
</table>

| Exclusion: | History of systemic inflammatory conditions; |
|           | Neurological deficits (motor, cognitive or sensory); |
|           | Recent lower limb injury or surgery (< 6 months); |
|           | History of other chronic pain disorders (e.g. fibromyalgia); |
|           | Skin allergies; |
|           | Allergy to menthol; |
|           | History of allergic reaction to other NSAIDs or aspirin; |
|           | Congestive heart failure (NYHA II-IV); |
|           | Inadequately controlled unstable hypertension (blood pressure persistently >140/90); |
|           | Ischaemic heart disease, peripheral arterial disease or cerebrovascular disease (including CABG surgery or angioplasty within the last year); |
|           | Severe hepatic dysfunction (serum albumin <25g/l); |
|           | Active GI bleeding or peptic ulceration; |
|           | Estimated creatinine clearance <30mL/min |
|           | Recent change (<1 month) in medications |
|           | Current use of high dose (>325mg daily) aspirin. |

All participants provided written informed consent before participating in the study. Ethical approval was provided in a single submission for studies 5, 6 and 7 by Royal
Perth Hospital Medical Research Ethics Committee (Approval EC 2009/100) and by Curtin University Human Research Ethics Committee (Approval number HR 26/2010).

OA participants using analgesics or non-steroidal anti-inflammatories (NSAIDs) were asked to withdraw from this medication for the course of the study. These subjects underwent a washout period equal to five half lives of their analgesic or NSAID. During this washout period participants were be able to use acetaminophen for rescue analgesia but were asked to refrain from its use 12 hours before testing.

Before baseline testing subjects were randomly assigned to either active or placebo group for the intervention phase of the study, using a computer generated algorithm administered by Royal Perth Hospital Pharmacy Department. All investigators were blind to group allocation.

In total, 86 OA subjects were recruited as six did not complete the study. Of these six, one withdrew before starting due to urgent need for joint injection, one withdrew before starting on the advice of his family doctor, and one participant withdrew due to inability to manage his pain. Three participants withdrew after 4 days due to non-serious adverse events. The results of the six subjects who withdrew were not included in the analysis. RPH Pharmacy allocated the six replacement OA subjects to Active or Placebo groups so that final group allocation was equal. These final 80 participants were those whose baseline results are reported in Chapter 8. Baseline characteristics are therefore shown in Table 8.3.

**Drug administration**

All drugs were administered by Royal Perth Hospital Pharmacy Department. This included all storage and blinded labelling of active and placebo drugs, creation and administration of the randomisation chart (to which all study investigators remained blind throughout the study), and dispensing of all study drugs, including rescue medication. All subjects were provided with rescue analgesia as part of their study drug pack, with the study Rheumatologist making the clinical decision about whether a participant required acetaminophen or tramadol. Only 5 participants were prescribed Tramadol. All study medication (active and placebo) was taken once per day. Subjects were advised to take rescue medication only if needed. Subjects were asked to refrain from taking their rescue medication before testing in order to wash out its potential effects: those taking Acetaminophen were asked to take their final dose 12 hours
before testing; those taking Tramadol took their final dose 24 hours before testing with additional acetaminophen if needed up to 12 hours before testing.

**Study Design and Procedures**

This intervention study used a double-blind, randomised, placebo-controlled design to compare 60mg daily etoricoxib with placebo formulation over a 14-day period. Subjects were tested at baseline when washed-out of their usual medication. This baseline data was reported and analysed in the cross-sectional study in Chapter 8. Following baseline testing subjects were provided with their active or placebo drug pack. They were then tested with the same outcome measures and procedures on Day 4 (steady plasma state) and on Day 14. Each test session was at the same time of day for each participant.

![Diagram](Diagram.png)

**Figure 9.1:** Etoricoxib intervention study protocol

All procedures were identical to those reported in Study 6 (Chapter 8). The baseline test session involved initial recording of self-reported co-morbidities and usual medications followed by completion of the questionnaires. Physical and quantitative sensory testing (QST) outcome measures were then completed. Day 0 test order was randomised between subjects, but within-subject test order was kept the same for subsequent sessions.

Following completion of informed consent procedures, subjects underwent a short interview to record any co-morbidities and usual medications. Subjects then completed
the four self-report questionnaires for quality of life, pain and function and presence of neuropathic pain (see below for details). All subjects then completed physical and quantitative sensory testing (QST) outcome measures described below. Order of physical and QST testing was randomised between modalities, although for punctate, heat and cold stimuli, detection threshold was always tested before pain threshold and the Aggregated Locomotion Function (ALF) tasks were always performed in the same order (chair, walk, stairs tasks).

**Outcome Measures**

All QST, physical and self-report outcome measures\(^1\) were the same as for Study 6 (Chapter 8) and shown below. Standardised method and instructions were used for each questionnaire and test. Although detection threshold were assessed as part of the study protocol, in part to ensure safety with pain threshold testing, they are not reported in the current study.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>QST</td>
<td>menthol-cold hyperalgesia • Algotype Descriptor Index (ADI):</td>
</tr>
<tr>
<td>(Cold detection)</td>
<td>1. VAS cold, heat, unpleasantness, pain;</td>
</tr>
<tr>
<td>Thermal-cold hyperalgesia</td>
<td>2. Pain quality descriptors (MWS)</td>
</tr>
<tr>
<td>(Heat detection)</td>
<td>Cold detection threshold (CDT) (°C)</td>
</tr>
<tr>
<td>Heat hyperalgesia</td>
<td>Cold pain threshold (CPT) (°C)</td>
</tr>
<tr>
<td>(Light touch detection)</td>
<td>Heat detection threshold (HDT) (°C)</td>
</tr>
<tr>
<td>Light touch allodynia</td>
<td>Heat pain threshold (HPT) (°C)</td>
</tr>
<tr>
<td>Mechanical hyperalgesia</td>
<td>Punctate detection threshold (PcDT)</td>
</tr>
<tr>
<td>Physical Function</td>
<td>Punctate pain threshold (PcPT)</td>
</tr>
<tr>
<td>Function</td>
<td>Pressure pain threshold (PPT) (kPa)</td>
</tr>
<tr>
<td>Aggregated Locomotion Function</td>
<td>2m return chair transfer Timed</td>
</tr>
<tr>
<td>Function score (ALF)</td>
<td>8m return walk (secs)</td>
</tr>
<tr>
<td></td>
<td>Stairs</td>
</tr>
<tr>
<td>Quality of Life</td>
<td>SF36-v2 Physical Health Score</td>
</tr>
<tr>
<td></td>
<td>Mental Health Score</td>
</tr>
</tbody>
</table>

\(^1\) All permissions granted for use of all questionnaires from copyright owners
All quantitative sensory tests were applied at the same standardised sites as for Studies 5 and 6 (Chapters 7 and 8). Menthol-cold hyperalgesia was tested at a 2x3cms site on the volar forearm 10cms above the wrist crease. All other QST tests were applied at the OA knee, the uninvolved knee and the ipsilateral elbow (Table 9.2). OA and ‘unaffected’ knees were designated according to highest VAS pain rating. All tests were applied using the identical equipment and standardised procedures as reported in Chapters 7 and 8. For each QST apart from the menthol test, one practice was always followed by 3 trials at each site, with the mean calculated for analysis.

<table>
<thead>
<tr>
<th>Pain and function</th>
<th>• Western Ontario &amp; Macmaster Universities (WOMAC) OA Knee Index: pain, stiffness, function subscores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuropathic pain</td>
<td>• PainDETECT</td>
</tr>
<tr>
<td>Pain quality</td>
<td>• Pain Quality Assessment Scale (PQAS)</td>
</tr>
</tbody>
</table>

| Table 9.2: Standardised knee and elbow test sites for QST tests |
|-----------------|------------------------------------------------------------------|
|                 | Punctate & Pressure Thresholds | Thermal Thresholds       |
| Knee            | Medial knee joint line                      | Distal quadriceps       |
| Elbow           | Common extensor origin                      | Distal triceps          |

For the menthol-cold test the same menthol Gel B from the same batch was used as for Study 6 (Chapter 8). The same e-VAS and descriptor list measurement system was also used, with ADI calculated as described previously. All procedures were identical.

**Sample Size and Statistical Analysis**

Power and sample size calculations were based on changes following intervention in PPT and in WOMAC score. CPT could not be considered since no previous studies have investigated change in CPT following intervention. For PPT, a 20% improvement from baseline mean following both manual and pharmaceutical interventions has been reported as clinically significance (Lemming et al., 2007; Moss et al., 2007). Studies investigating Cox-2 interventions have reported clinically significant change in WOMAC score of 20-22% (Bingham et al., 2007). With alpha set at p<0.05, it was calculated that an equal sample size of between 37 and 45 per group would provide 80% power to
detect a 20% difference in both PPT and WOMAC between experimental and placebo
groups at the 14-day time-point. This calculation was based on actual between-group
difference for PPT of 38kPa (standard deviation 57kPa) (Moss et al., 2007) and an
actual difference of 10.1mm and standard deviation of 17mm in WOMAC score
(Bingham et al., 2007).

Normality testing was carried out using a combination of Shapiro-Wilk tests and visual
analysis of distribution, as described in Chapter 8 for baseline testing. All but CPT data
was normally distributed so that parametric statistics could be applied. For ADI
distribution at baseline, see Figure 8.2. For CPT, instead of a single Repeated Measures
ANOVA with between-subject factor for Active and Placebo groups, Friedman’s two-
way ANOVA by ranks had to be applied to Active and Placebo data individually.

The following data analyses were carried out using SPSS statistics package, version 19
with alpha set at p<.05.

<table>
<thead>
<tr>
<th>Hypotheses and Data</th>
<th>Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. a) Significant improvements in self-reported pain, dysfunction and quality of life by Day 14 for those taking Active etoricoxib:</td>
<td>Repeated Measures ANOVA (Day 0, 4, 14); between-subjects factor - Active / Placebo group</td>
</tr>
<tr>
<td>• WOMAC-pain sub-score</td>
<td></td>
</tr>
<tr>
<td>• Mean VAS pain in last 7 days</td>
<td></td>
</tr>
<tr>
<td>• WOMAC function sub-score</td>
<td></td>
</tr>
<tr>
<td>• ALF: chair time; walk time; stairs time (secs)</td>
<td></td>
</tr>
<tr>
<td>• SF-36 PCS score</td>
<td></td>
</tr>
<tr>
<td>b) Significant reductions in measures of local hyperalgesia or inflammation by Day 14 for those taking Active etoricoxib:</td>
<td>Friedman’s two-way ANOVA by ranks or Repeated Measures ANOVA (Day 0, 4, 14); between-subjects factor - Active / Placebo group</td>
</tr>
<tr>
<td>• OA knee CPT (°C)</td>
<td></td>
</tr>
<tr>
<td>• OA knee PPT (kPa)</td>
<td></td>
</tr>
<tr>
<td>• OA knee HPT (°C)</td>
<td></td>
</tr>
<tr>
<td>• OA knee PcPT (log)</td>
<td></td>
</tr>
<tr>
<td>• PQAS – deep sub-scale</td>
<td></td>
</tr>
<tr>
<td>c) Significant reductions in widespread hyperalgesia measures by Day 14 for those</td>
<td></td>
</tr>
<tr>
<td>Taking Active etoricoxib:</td>
<td>Friedman's two-way ANOVA by ranks or Repeated Measures ANOVA (Day 0, 4, 14); between-subjects factor - Active / Placebo group</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>• ADI score</td>
<td></td>
</tr>
<tr>
<td>• CPT (°C) unaffected knee &amp; elbow</td>
<td></td>
</tr>
<tr>
<td>• PPT (kPa) unaffected knee &amp; elbow</td>
<td></td>
</tr>
<tr>
<td>• HPT (°C) unaffected knee &amp; elbow</td>
<td></td>
</tr>
<tr>
<td>• PcPT (log) unaffected knee &amp; elbow</td>
<td></td>
</tr>
<tr>
<td>d) Significant improvements in measures of neuropathic-type pain measures by Day 14 for those taking Active etoricoxib:</td>
<td>Repeated Measures ANOVA (Day 0, 4, 14); between-subjects factor - Active / Placebo group</td>
</tr>
<tr>
<td>• PainDETECT score</td>
<td></td>
</tr>
<tr>
<td>• PQAS sub-scores</td>
<td></td>
</tr>
<tr>
<td>2. Measures of cold hyperalgesia (mean CPT and ADI) will show similar responsiveness to etoricoxib over 14 days:</td>
<td>Effect size analysis Number needed to treat analysis</td>
</tr>
<tr>
<td>• Effect size</td>
<td></td>
</tr>
<tr>
<td>• NNT</td>
<td></td>
</tr>
<tr>
<td>3. Significant differences in response to etoricoxib between those with and without menthol-cold hyperalgesia (ADI &lt; or ≥5):</td>
<td>Comparison of separate analyses for ADI &lt;5 and ADI ≥5:</td>
</tr>
<tr>
<td>• WOMAC pain, stiffness &amp; function</td>
<td>Repeated Measures ANOVA (Day 0, 4, 14); between-subjects factor - Active / Placebo group</td>
</tr>
<tr>
<td>• Mean VAS pain (last 7 days)</td>
<td></td>
</tr>
<tr>
<td>• ADI, CPT, PPT, HPT</td>
<td></td>
</tr>
<tr>
<td>• PainDETECT &amp; PQAS paroxysmal, surface and deep sub-scores</td>
<td></td>
</tr>
<tr>
<td>• AFL score: chair, walk, stairs tasks</td>
<td></td>
</tr>
<tr>
<td>• SF-36 PCS scores</td>
<td></td>
</tr>
</tbody>
</table>
9.5 Results

Baseline Characteristics

- There were no significant differences between Placebo and Active groups in any baseline characteristics

The total OA cohort comprised 36 male and 44 female subjects with a mean age of 64 years (range 50-86 years). When divided into Placebo and Active groups, this balance remained, with no significant differences in gender between groups (p=.822). Similarly there was no significant difference between Placebo and Active groups for age or Body Mass Index (Table 9.3).

Table 9.3: Baseline characteristics of Placebo and Active groups

<table>
<thead>
<tr>
<th></th>
<th>Placebo (n=40)</th>
<th>Active (n=40)</th>
<th>t(78)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (M:F)</td>
<td>19:21</td>
<td>17:23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (range) years</td>
<td>65 (50-86)</td>
<td>65 (50-80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (SD)</td>
<td>29.3 (4.5)</td>
<td>27.8 (4.7)</td>
<td>.146</td>
<td>.148</td>
</tr>
</tbody>
</table>

Baseline self-reported medications and co-morbidities are reported in greater detail in Chapter 8. The majority of subjects reported regular use of either slow release high dose acetaminophen (‘Panadol-Osteo’) or over-the-counter non-steroidal anti-inflammatory drugs (NSAIDs) for analgesia. The most frequently reported medicated comorbidities were diabetes and high blood pressure (both for 29% of the cohort).

- There were no significant differences between Placebo and Active groups for baseline self-reported pain, dysfunction or quality of life.

There were no significant differences between Placebo and Active groups for self-reported pain, dysfunction or quality of life measures (Table A4.1 in Appendix 4). On average the whole OA cohort reported moderate pain and functional disability. Average VAS pain in the last 7 days was 4.6/10 (range 3-8/10). There were moderate to high levels of perceived stiffness (WOMAC stiffness: mean 8.2/20; max 18.9/20) and moderate levels of functional disability (WOMAC function: mean 53.4/250; max 123/170). Both groups of subjects reported well below average quality of life related to physical functioning (PCS) but slightly above average scores for quality of life related to mental health (MCS).

- There were no significant differences between Placebo and Active groups for baseline PainDETECT or PQAS sub scores.
Active and Placebo groups also reported similar levels of neuropathic-type pain symptoms. There was no significant group difference in mean PainDETECT scores at baseline and the 13.8% of subjects (n=11) who scored as positive for neuropathic pain in total were split evenly between groups. No significant differences were seen for PQAS scores at baseline. (Table A4.1 in Appendix 4).

- There were also no differences in baseline physical function tests and only minimal differences between Placebo and Active groups for baseline QST tests.

The only group differences at baseline were seen for HPT at the elbow (p=.038) and for punctate pain threshold at the unaffected knee (p=.016) (Table A4.2 in Appendix 4).

**Hypothesis 1a**

*Subjects with knee OA who receive a daily dose of etoricoxib 60mg for 14 days will demonstrate a significant improvement in pain, dysfunction and quality of life when compared with subjects receiving placebo formulation.*

- Those in the Active group showed significantly greater reduction in WOMAC and VAS pain scores.

There was a significant group x time interaction effect for WOMAC-pain scores (F(2,156) = 10.85, p<.001). Active group subjects reported a mean 30.6% decrease in pain from Day 0 to Day 14, compared with a 10% increase in pain for the Placebo group (Figure 9.2). There was also a significant group x time interaction effect for VAS pain (last 7 days) (F(2,156)= 3.16, p<.001). Active participants’ pain reduced on average from 4.6/10 to 3.6/10 (22% reduction) whereas Placebo participants’ pain increased from 4.6/10 to 4.9/10 (7% increase).

![Figure 9.2: Mean WOMAC-pain scores for Placebo and Active groups at each time point.](image)

- Those in the Active group significantly improved their WOMAC function and stiffness scores, compared with Placebo.
There was a significant group x time interaction effect for WOMAC-function scores ($F_{(2, 156)} = 15.21$, $p<.001$), with the Active group improving their scores by $28.4\%$ by Day 14. In contrast, the function scores of Placebo group subjects reduced by $14\%$.

![Figure 9.3: Mean WOMAC-function scores for Placebo and Active groups at each time point.](image)

WOMAC-stiffness also showed a significant group x time interaction effect ($F_{(2, 156)} = 15.75$, $p<.001$). Active participants showed an improvement of $25.6\%$ by Day 3 and $44\%$ by Day 14, in contrast to the $9\%$ worsening in stiffness reported by the Placebo group.

- **Active group participants showed a significant improvement in time taken to complete all functional (ALF) tasks and in pain experienced during the tasks**

There was a statistically significant group x time interaction effect for each ALF task (Table 9.4). The greatest difference was shown for the stairs task where Active group participants decreased their time by $13.2\%$, in contrast to the Placebo group whose time increased marginally by $3.4\%$.

<table>
<thead>
<tr>
<th></th>
<th>Day 0</th>
<th>Day 4</th>
<th>Day 14</th>
<th>$F_{(2, 156)}$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>6.4 (.31)</td>
<td>6.3 (.31)</td>
<td>6.3 (.33)</td>
<td>7.77</td>
<td>.003*</td>
</tr>
<tr>
<td>Active</td>
<td>6.9 (.36)</td>
<td>6.2 (.26)</td>
<td>6.0 (.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>12.6 (.53)</td>
<td>12.4 (.47)</td>
<td>13.0 (.63)</td>
<td>5.38</td>
<td>.006*</td>
</tr>
<tr>
<td>Active</td>
<td>13.2 (.61)</td>
<td>12.5 (.57)</td>
<td>12.4 (.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stairs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>17.3 (1.4)</td>
<td>17.7 (1.4)</td>
<td>17.9 (1.4)</td>
<td>9.49</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>Active</td>
<td>19.0 (1.6)</td>
<td>17.0 (1.4)</td>
<td>16.5 (1.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

VAS pain ratings before the start of the tasks, during each task and after each task were also analysed. Resting pain at the start was less for both groups on Day 14: $69.2\%$ less for Active and $35\%$ less for Placebo. Although pain ratings were less for all tasks for
both groups on Day 14, the overall improvement in pain from start to completion was significantly greater for Active participants ($F(1, 156) = 5.94$, $p = .017$). The pain during the stairs task was the most discriminatory: pain increased by 21% less for Active participants by Day 14 whereas it more than doubled for Placebo participants.

**Figure 9.4:** VAS pain rating before, during and after ALF chair, walk and stairs tasks on Day 0 and Day 14: a) Active group; b) Placebo group

- The Active group also showed significant improvements in quality of life related to both physical function and mental health

There was a significant group x time interaction effect for both SF-36 PCS sub-score ($t(79) = -4.55$, $p < .001$) and MCS sub-score ($t(79) = -2.70$, $p = .010$). Active participants improved their physical score by 13% and their mental health score by 7%, whereas Placebo participants’ scores reduced by 4-5%.

**Hypothesis 1b**

*Knee OA subjects receiving etoricoxib 60mg for 14 days will demonstrate a significant reduction in local hyperalgesia when compared with subjects receiving placebo formulation.

- Cold pain threshold temperature showed a statistically significant improvement at the OA knee by Day 14 for the group taking Active etoricoxib

CPT at the OA knee improved (decreased) by approximately 10% for the Active group and increased by 12% for the Placebo group (Figure 9.5). Friedman's Two-Way ANOVA analyses showed that the change in CPT by Day 14 was significant for the Active group ($\chi^2(2) = 6.61$, $p = .037$) but not for the Placebo group ($\chi^2(2) = .053$, $p = .974$).
Pressure pain threshold and punctate pain values also significantly improved for the Active group at the OA knee by Day 14.
There was also a significant group x time interaction effect for PPT at the OA knee ($F_{1.56}$= 9.67 , $p<.001$). This is shown with all other PPT results in Table 9.6. This amounted to a 32% improvement in PPT at the OA knee compared with the 2% decline in PPT for the Placebo group at the OA knee. Punctate pain threshold also showed a significant group x time interaction effect at the OA knee ($F_{1.56}$= 11.48 , $p<.001$).

There was no difference in change between Active and Placebo groups for heat pain threshold at the OA knee.
HPT showed no significant group x time interaction effect (Table 9.8): OA knee: $F_{1.56}$= 1.72 , $p=.183$.

**Hypothesis 1c**

Knee OA subjects receiving etoricoxib 60mg for 14 days will demonstrate a significant reduction in widespread hyperalgesia when compared with subjects receiving placebo formulation.

By Day 14 subjects in the Active group exhibited significant improvements in forearm ADL, a measure of widespread cold response.
There was a significant group x time interaction effect for ADL score ($F_{1.56}$= 7.17 , $p<.001$) (Figure 9.6) with the Active group reducing their score by 20% and Placebo group increasing by 13%.
Break-down of the ADI into component parts showed that both sensation quality and intensity changed significantly for the Active group by Day 14 (Table 9.5). Closer examination of word choice showed that change in MWS score was driven by change in selection of the words burning and prickling: burning was selected by 10% fewer Active participants by Day 14 but by 15% more of the Placebo group; choice of prickling also reduced by 10% for the Active group but remained unchanged for the Placebo group. VAS intensity values for unpleasantness and pain were significantly reduced for the Active group and increased for the Placebo group. VAS cold reduced significantly for the Active group but remained unchanged for the Placebo group. Heat VAS did not significantly change by Day 14 for either Active or Placebo groups (Table 9.5).

**Table 9.5:** Change in ADI score components: mean (SEM) values for sensation quality (MWS) and intensity (VAS cold, heat, unpleasantness and pain) for Active and Placebo groups at Days 0, 4 and 14.

<table>
<thead>
<tr>
<th>ADI score component</th>
<th>Day 0</th>
<th>Day 4</th>
<th>Day 14</th>
<th>F(1,156) or ( \chi^2(2) )</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MWS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>2.4 (.14)</td>
<td>2.5 (.13)</td>
<td>2.6 (.14)</td>
<td>7.70</td>
<td>.001**</td>
</tr>
<tr>
<td>Active</td>
<td>2.6 (.12)</td>
<td>2.3 (.12)</td>
<td>2.3 (.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VAS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cold</td>
<td>42.9 (4.4)</td>
<td>42.5 (4.3)</td>
<td>40.2 (4.5)</td>
<td>.419</td>
<td>.811</td>
</tr>
<tr>
<td>Placebo</td>
<td>35.5 (3.3)</td>
<td>30.6 (3.2)</td>
<td>28.3 (3.0)</td>
<td>8.01</td>
<td>.018*</td>
</tr>
<tr>
<td>Active</td>
<td>5.1 (2.1)</td>
<td>11.1 (3.3)</td>
<td>9.5 (2.1)</td>
<td>6.14</td>
<td>.055</td>
</tr>
<tr>
<td>Heat</td>
<td>4.5 (1.7)</td>
<td>4.4 (1.7)</td>
<td>3.2 (1.4)</td>
<td>.211</td>
<td>.900</td>
</tr>
<tr>
<td>Placebo</td>
<td>12.6 (3.2)</td>
<td>16.9 (3.7)</td>
<td>18.5 (4.1)</td>
<td>6.23</td>
<td>.044*</td>
</tr>
<tr>
<td>Active</td>
<td>16.6 (3.7)</td>
<td>9.6 (2.7)</td>
<td>9.2 (2.8)</td>
<td>11.78</td>
<td>.003*</td>
</tr>
<tr>
<td>Unpleasantness</td>
<td>4.4 (2.3)</td>
<td>4.7 (2.5)</td>
<td>7.4 (3.0)</td>
<td>6.07</td>
<td>.048*</td>
</tr>
<tr>
<td>Placebo</td>
<td>2.3 (1.2)</td>
<td>1.8 (1.1)</td>
<td>.93 (0.7)</td>
<td>5.77</td>
<td>.050*</td>
</tr>
</tbody>
</table>
• **Subjects in the Active group did not show widespread reduction in CPT.**

There was no significant effect for CPT by Day 14 at the unaffected knee or at the elbow. Friedman’s ANOVA analyses showed no significant change in CPT at the unaffected knee for the Active group ($\chi^2(2) = 1.75, p = .417$) or for the Placebo group ($\chi^2(2) = 3.07, p = .215$). There was no significant change at the elbow either: Active group $\chi^2(2) = 1.50, p = .472$; Placebo group ($\chi^2(2) = 5.59, p = .061$). Despite the very small mean changes in temperature, there was wide inter-subject variation for both groups. For example, at the unaffected knee site Active group CPT values ranged from 10.2°C improvement to 17.8°C decline, whereas the Placebo group ranged from 15.2°C improvement to 22.5°C decline.

![Figure 9.7a-b: Mean (SEM) cold pain threshold values for Placebo and Active groups at each time point for a) unaffected knee; b) elbow.](image-url)

• **By Day 14 Active group participants exhibited a significant improvement in PPT at unaffected knee but not at the elbow (Table 9.6).**

PPT at the unaffected knee showed significant group x time interaction effects (unaffected knee: $F_{(1,156)} = 7.96, p = .001$) but there was no significant interaction effect at the elbow ($F_{(1,156)} = .942, p = .386$). PPT at the unaffected knee improved by 14% from Day 0 to Day 14 for the Active group, whereas PPT for the Placebo group declined by 7%. Both Active and Placebo groups increased their elbow PPT: Active group by 9.5%, Placebo group by 4.7%.
Table 9.6: Mean (SEM) values pressure pain thresholds at the OA knee, unaffected knee and elbow test sites for Active and Placebo groups on Days 0, 4 and 14.

<table>
<thead>
<tr>
<th></th>
<th>Day 0</th>
<th>Day 4</th>
<th>Day 14</th>
<th>F(1, 156)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PPT</strong></td>
<td>OA knee</td>
<td>Placebo</td>
<td>Active</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>268.3(21)</td>
<td>267.4(21)</td>
<td>272.9(20)</td>
<td>9.67</td>
<td>.001**</td>
</tr>
<tr>
<td></td>
<td>283.2(18)</td>
<td>347.1(20)</td>
<td>372.7(25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>Placebo</td>
<td>325.8(23)</td>
<td>267.4(21)</td>
<td>302.1(21)</td>
<td>7.96</td>
</tr>
<tr>
<td></td>
<td>330.6(20)</td>
<td>347.1(20)</td>
<td>376.2(27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elbow</td>
<td>Placebo</td>
<td>331.5(20)</td>
<td>325.1(21)</td>
<td>347.0(19)</td>
<td>.942</td>
</tr>
<tr>
<td></td>
<td>350.1(15)</td>
<td>369.3(16)</td>
<td>383.4(16)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Active participants also showed a significant increase in punctate pain threshold at the unaffected knee but not at the elbow.**

PcPPT at the unaffected knee showed significant group x time interaction effects (unaffected knee: F(1, 156)= 6.25 , p = .004) but there was no significant interaction effect at the elbow (F(1, 156)= 2.65 , p = .074). At both unaffected knee and elbow, punctate pain sensitivity decreased for the Active group whilst remaining unchanged for the Placebo group.

- **There was no significant change in HPT temperatures at any of the test sites.**

Mean HPT showed minimal change for either group at any of the test sites (Table 9.7).

Table 9.7: Mean (SEM) values heat pain thresholds at the OA knee, unaffected knee and elbow test sites for Active and Placebo groups on Days 0, 4 and 14.

<table>
<thead>
<tr>
<th></th>
<th>Day 0</th>
<th>Day 4</th>
<th>Day 14</th>
<th>F(1, 76)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HPT</strong></td>
<td>OA knee</td>
<td>Placebo</td>
<td>Active</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44.8 (.48)</td>
<td>45.7 (.54)</td>
<td>45.9 (.45)</td>
<td>1.72</td>
<td>.183</td>
</tr>
<tr>
<td></td>
<td>45.8 (.49)</td>
<td>45.6 (.40)</td>
<td>46.1 (.42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>Placebo</td>
<td>45.2 (.50)</td>
<td>45.4 (.54)</td>
<td>46.2 (.43)</td>
<td>.277</td>
</tr>
<tr>
<td></td>
<td>45.8 (.44)</td>
<td>45.7 (.41)</td>
<td>46.7 (.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elbow</td>
<td>Placebo</td>
<td>45.5 (.65)</td>
<td>46.0 (.59)</td>
<td>46.1 (.54)</td>
<td>.618</td>
</tr>
<tr>
<td></td>
<td>47.3 (.54)</td>
<td>47.2 (.49)</td>
<td>46.7 (.34)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Hypothesis 1d**

*Knee OA subjects receiving etoricoxib 60mg for 14 days will demonstrate significant improvements in neuropathic-type pain scores when compared with subjects receiving placebo formulation.*
• **By Day 14 Active group participants reported a significant reduction in PainDETECT score.**

There was a significant group x time interaction effect for PainDETECT ($F_{(1, 76)}= 19.04, p<.001$) (Figure 9.8). Active PainDETECT scores decreased steadily from 12.0 ($\pm .84$) to 9.2 ($\pm .9$) whereas Placebo scores increased from 11.9 ($\pm .83$) to 13.6 ($\pm 1.0$).

![Figure 9.8: Mean (SEM) PainDETECT values for Placebo and Active groups at each time point](image)

This change in total score translated into significant changes in PainDETECT category between Day 0 and Day 14 for Placebo and Active groups (Table 9.8). 10 Placebo group subjects and 12 Active subjects moved category. 7 of the 10 Placebo group subjects moved to a more severe category: 4 from ‘unclear’ to ‘positive neuropathic’ and 3 from ‘negative’ to ‘unclear’. 3 moved from ‘unclear’ to ‘negative’. In contrast, 11 of the 12 Active subjects moved to a less severe category: 3 from ‘positive neuropathic’ to ‘negative neuropathic’, 1 from ‘positive’ to ‘unclear’ and 7 from ‘unclear’ to ‘negative’.

**Table 9.8:** Change in PainDETECT category (‘negative neuropathic’ <13; ‘unclear neuropathic’ 13-18; ‘positive neuropathic’ ≥19) between Day 0 and Day 14 for Placebo and Active groups

<table>
<thead>
<tr>
<th>Placebo Group</th>
<th>Day 14</th>
<th></th>
<th></th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;13</td>
<td>13-18</td>
<td>≥19</td>
<td></td>
</tr>
<tr>
<td>Day 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;13</td>
<td>17</td>
<td>3</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>≥19</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>total</td>
<td>19</td>
<td>12</td>
<td>9</td>
<td>40</td>
</tr>
</tbody>
</table>

χ² = 37.5, p < .001

<table>
<thead>
<tr>
<th>Active Group</th>
<th>Day 14</th>
<th></th>
<th></th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;13</td>
<td>13-18</td>
<td>≥19</td>
<td></td>
</tr>
<tr>
<td>Day 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;13</td>
<td>22</td>
<td>1</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>13-18</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>≥19</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>total</td>
<td>32</td>
<td>5</td>
<td>3</td>
<td>40</td>
</tr>
</tbody>
</table>

χ² = 19.9, p = .001
• All PQAS sub-scores were significantly reduced by Day 14 in the Active group, although the greatest improvement was seen for the ‘deep’ sub-score.

Each sub-score of the PQAS questionnaire also showed a significant group x time interaction effect: PQAS paroxysmal $F_{(1,76)} = 17.77, p<.001$; surface $12.19, p<.001$; deep $24.58, p<.001$ (Figure 9.9).

![Figure 9.9: Mean (SEM) PQAS sub-scores for Placebo and Active groups at each time point: paroxysmal, surface and deep sub-scores.](image)

**Hypothesis 2**

The measures of cold pain threshold and ADI will show similar responsiveness to the intervention of etoricoxib 60mg for 14 days (Table 9.9).

• Percentage change from baseline by Day 14 was not the same: ADI percentage change was higher than for CPT and similar to that for PainDETECT

Table 9.9 shows that basic percentage change from baseline (Day 0) to Day 14 was greater for ADI than for CPT mean and CPT at the OA knee. ADI improved overall by 20.5% in Active participants with Placebo participants worsening by 13.5%. This pattern was very similar for the neuropathic pain questionnaire PainDETECT. Mean CPT however only improved by a minimal 4% and OA knee CPT by 8.9%. All both cases, Placebo CPT reduced significantly, resulting in a net change of around 20%.

• Cohen’s $d$ and effect size were not the same: Cohen’s $d$ and effect size was higher for ADI and PainDETECT than for CPT

Cohen’s $d$ and effect size coefficient were also calculated by comparing mean (SD) change from Day 0 to Day 14 between Active and Placebo groups. A similar pattern was seen, with both values considerably greater for ADI and PainDETECT than for either mean CPT or CPT at the OA knee (Table 9.9). Effect sizes for ADI and PainDETECT were moderate whereas those for mean and OA knee CPT values were small.
• **Numbers needed to treat (NNT) analyses were not the same:** ADI and PainDETECT showed greater responsiveness to etoricoxib than CPT (Table 9.9).

NNT analysis comparison showed that for a 30% reduction in ADI score or PainDETECT score 2.9 and 2.2 were needed. For a 30% reduction in CPT temperature, 5 were needed for a significant mean change and 8 for change at the OA knee alone.

**Table 9.9:** Comparison of different measures of responsiveness to etoricoxib treatment for ADI, PainDETECT, cold pain threshold and pressure pain threshold.

<table>
<thead>
<tr>
<th></th>
<th>% Change in value Day 0-14</th>
<th>Cohen’s d</th>
<th>Effect size (r)</th>
<th>Numbers Needed to Treat (NNT)(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>+ 20.5%</td>
<td>1.64</td>
<td>.633</td>
<td>30% reduced 2.9</td>
</tr>
<tr>
<td>Placebo</td>
<td>- 13.5%</td>
<td></td>
<td></td>
<td>50% reduced 6.7</td>
</tr>
<tr>
<td><strong>PainDETECT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>+ 23%</td>
<td>1.47</td>
<td>.593</td>
<td>30% reduced 2.2</td>
</tr>
<tr>
<td>Placebo</td>
<td>- 14%</td>
<td></td>
<td></td>
<td>50% reduced 4</td>
</tr>
<tr>
<td><strong>CPT mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>+ 4%</td>
<td>.591</td>
<td>.283</td>
<td>30% reduced 5</td>
</tr>
<tr>
<td>Placebo</td>
<td>- 16%</td>
<td></td>
<td></td>
<td>50% reduced 20</td>
</tr>
<tr>
<td><strong>OA knee</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>+ 8.9%</td>
<td>.507</td>
<td>.246</td>
<td>30% reduced 8</td>
</tr>
<tr>
<td>Placebo</td>
<td>- 12%</td>
<td></td>
<td></td>
<td>50% reduced 10</td>
</tr>
<tr>
<td><strong>PPT mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>+17%</td>
<td>.680</td>
<td>.322</td>
<td>30% improved 40</td>
</tr>
<tr>
<td>Placebo</td>
<td>-.40%</td>
<td></td>
<td></td>
<td>50% improved 20</td>
</tr>
<tr>
<td><strong>OA knee</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>+32%</td>
<td>.260</td>
<td>.129</td>
<td>30% improved 4</td>
</tr>
<tr>
<td>Placebo</td>
<td>-2%</td>
<td></td>
<td></td>
<td>50% improved 5.7</td>
</tr>
</tbody>
</table>

\(^1\) NNT analysis calculated as \(1/(\text{No of outcomes Active group})-(\text{No of outcomes Placebo group})\). As shown, both 30% reduction by Day 14 and 50% reduction by Day 14 were calculated.

**Hypothesis 3**

**Subjects with cold hyperalgesia (ADI ≥5) will respond less well to etoricoxib than those with a less hyperalgesic response (ADI<5).**

Subjects were separated post hoc into those with menthol ADI score < or ≥5 and differences between Placebo and Active for change over time analysed separately. Table 9.10 shows the number of subjects in each ADI group and their Placebo and Active drug allocations. Group comparisons for Day 0 are shown in Chapter 8.
Table 9.10: Numbers of subjects with ADI < or ≥ 5 allocated to Placebo or Active groups

<table>
<thead>
<tr>
<th></th>
<th>Placebo (n)</th>
<th>Active (n)</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low ADI (&lt;5)</td>
<td>26</td>
<td>29</td>
<td>55</td>
</tr>
<tr>
<td>High ADI (≥5)</td>
<td>14</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td><strong>40</strong></td>
<td><strong>40</strong></td>
<td><strong>80</strong></td>
</tr>
</tbody>
</table>

- Subjects with menthol ADI≥5 showed less clear improvements in pain, function and quality of life, compared with those in the ADI<5 group.

There were significant group x time interaction effects for pain VAS and all WOMAC sub-scores for both ADI groups (Figure 9.10a-h). The interactions for participants with ADI ≥5 were slightly less strong than for those with ADI<5.
**Figure 9.10a-h:** Mean (SEM) Active and Placebo group pain VAS scores at Day 0, 4 and 14 for low and high ADI groups.

There were significant group x time interaction effects for both ADI groups for ALF chair and stairs tasks (Fig 9.11a) and c). The walk task showed a significant interaction effect for the low ADI group only (Fig 9.11c).

**Figure 9.11a-c:** Mean (SEM) for ALF physical function tasks at Day 0, 4 and 14 for low and high ADI groups: a) chair task; b) walk task; c) stairs task.
SF-36 quality of life physical function scores also showed significant group x time interactions, although the effects were greater for the low ADI group: ADI<5 $F_{(1,53)}=20.17, p<.001$; ADI ≥ 5 $F_{(1,53)}=5.72, p=.025$.

- **Menthol ADI groups improved similarly in total PainDETECT score by Day 14.**

There were significant group x time interaction effects for PainDETECT scores for both ADI groups (Figure 9.12a-b). Although actual scores were significantly higher at all time points for those with ADI ≥5, the change interaction was similar between ADI groups (ADI<5: Active 22% decrease by Day 14, Placebo 15% increase; ADI≥5: Active 25% decrease, Placebo 13% increase). There was a slight group difference in PainDETECT classification change. At baseline, 80% of those in the high ADI group and 31% of those in the low ADI group scored as ‘unclear’ or ‘positive neuropathic’. By Day 14, there was little classification change for the low ADI participants, but for the high ADI group nearly 30% had reduced their scores sufficiently to be classified as ‘negative neuropathic’ (Table 9.11).

![Graphs showing PainDETECT scores for ADI<5 and ADI≥5 groups](image)

**Figure 9.12a-b:** Mean (SEM) PainDETECT scores at Day 0, 4 and 14 for low and high ADI groups: a) ADI<5 group; b) ADI≥5 group.

| Table 9.11: PainDETECT group change by ADI group, Day 0 to Day 14 |
|----------------|---------------|---------------|---------------|---------------|
|                | Day 0         | Day 14        |                |               |
|                | ADI<5 n=55    | ADI≥5 n=25    | ADI<5 n=55    | ADI≥5 n=25    |
| -ve neuropathic (<13) | 38 (69%)      | 39 (71%)      | 12 (48%)      |                |
| Unclear (13-18)     | 13 (24%)      | 12 (48%)      | 5 (20%)       | 12 (48%)      |
| +ve neuropathic (≥19) | 4 (7%)        | 8 (32%)       | 4 (7%)        | 8 (32%)       |

- **For PQAS, those with low ADI made greater improvements in ‘deep’ and ‘paroxysmal’ sub-scores than those with high ADI.**

291
There were also significant group x time interaction effects for all PQAS sub-scores for both high and low ADI groups, although some differences in magnitude of response were seen. Active participants with low ADI showed a 50% reduction in ‘deep’ sub-score sensations (throbbing, aching, tender) (Figure 9.13a), and a 45% reduction in ‘paroxysmal sensations (shooting, sharp, radiating), compared with 30% reduction for both sub-scores for the high ADI-Active group (Figure 9.13b). Change for ‘surface’ score was proportionately similar.

**Figure 9.13a:** Mean (SEM) PQAS ‘paroxysmal’, ‘surface’ and ‘deep’ sub-scores at Day 0, 4 and 14 for the low ADI group.

**Figure 9.13b:** Mean (SEM) PQAS ‘paroxysmal’, ‘surface’ and ‘deep’ sub-scores at Day 0, 4 and 14 for the high ADI group.

- Cold hyperalgesic and non-hyperalgesic participants showed similar reductions in their menthol ADI score and similar lack of change in their cold pain thresholds.
There were significant group x time interaction effects for menthol ADI scores for both ADI groups (Figure 9.14a-b). For both groups, Active participants reduced their ADI score by around 20%, while Placebo subjects increased their scores.

**Figure 9.14a-b:** Mean (SEM) menthol ADI scores at Day 0, 4 and 14 for a) the low ADI group and b) the high ADI group.

In contrast, there was minimal change in CPT for either ADI group (Table 9.12).

**Table 9.12:** Mean (SEM) cold pain threshold values at Day 0, 4 and 14 for low and high ADI groups.

<table>
<thead>
<tr>
<th>CPT</th>
<th>ADI&lt;5</th>
<th>ADI≥5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo</td>
<td>Active</td>
</tr>
<tr>
<td>OA knee</td>
<td>0</td>
<td>10.4(2.0)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>11.2(2.0)</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>11.0(1.9)</td>
</tr>
<tr>
<td></td>
<td>X²(2)= 3.06</td>
<td>X²(2)= 6.39</td>
</tr>
<tr>
<td></td>
<td>p=.216</td>
<td>p=.041*</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>0</td>
<td>6.5(1.6)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8.9(1.8)</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>10.1(1.9)</td>
</tr>
<tr>
<td></td>
<td>X²(2)= 4.07</td>
<td>X²(2)= 1.45</td>
</tr>
<tr>
<td></td>
<td>p=.131</td>
<td>p=.485</td>
</tr>
<tr>
<td>Elbow</td>
<td>0</td>
<td>6.4(1.4)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7.0(1.4)</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>7.6(1.4)</td>
</tr>
<tr>
<td></td>
<td>X²(2)= 1.20</td>
<td>X²(2)= 2.48</td>
</tr>
<tr>
<td></td>
<td>p=.550</td>
<td>p=.289</td>
</tr>
</tbody>
</table>

- There were differences between cold hyperalgesic groups in PPT change, with Active participants with low ADI showing a significant improvement compared with Placebo.
Mechanical cold hyperalgesia (PPT) showed response differences between ADI groups (Figure 9.15a-b). Participants with ADI<5 showed a significant group x time interaction effect for PPT at the OA knee (p=.001) and also at the unaffected knee (p=.001) but not at the elbow (p=.471). The only significant interaction effect for those with ADI ≥5 was for the OA knee (p=.038).

![Figure 9.15a](image)

**Figure 9.15a**: Mean (SEM) pressure pain threshold values at Day 0, 4 and 14 for the low ADI group for OA knee, unaffected knee and elbow test sites.

![Figure 9.15b](image)

**Figure 9.15b**: Mean (SEM) pressure pain threshold values at Day 0, 4 and 14 for the high ADI group for OA knee, unaffected knee and elbow test sites.

- Effect size calculations showed similarities in response between low and high ADI groups, although there was a trend towards those with low ADI responding slightly better on PPT and some measures of physical function (Table 9.13).

Cohen’s d and effect size correlations were calculated for each ADI group, comparing mean (SD) change for Day 0 and Day 14 between Active and Placebo groups. Overall there was little clear difference in response between those with low and high menthol ADI scores. There was a slight trend towards Active participants in the low ADI group
responding slightly better as far as mechanical hyperalgesia and the chair transfer and walk tests of physical function were concerned.

**Table 9.13**: Comparison between ADI groups of effect size and Cohen’s d for key study outcome measures.

<table>
<thead>
<tr>
<th></th>
<th>ADI &lt;5 Cohen’s d</th>
<th>ADI ≥5 Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect size (r)</td>
<td>Effect size (r)</td>
</tr>
<tr>
<td>ADI</td>
<td>1.03</td>
<td>9.92</td>
</tr>
<tr>
<td>PainDETECT</td>
<td>1.07</td>
<td>1.43</td>
</tr>
<tr>
<td>PPT</td>
<td>mean 0.881</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>OA knee 0.934</td>
<td>0.460</td>
</tr>
<tr>
<td>CPT</td>
<td>mean 0.481</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>OA knee 0.402</td>
<td>0.809</td>
</tr>
<tr>
<td>WOMAC</td>
<td>pain 0.874</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td>stiffness 1.17</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>function 1.03</td>
<td>1.16</td>
</tr>
<tr>
<td>ALF</td>
<td>chair 0.542</td>
<td>0.469</td>
</tr>
<tr>
<td></td>
<td>walk 0.922</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td>stairs 0.528</td>
<td>1.24</td>
</tr>
</tbody>
</table>

An additional analysis calculated effect size and Cohen’s d values for the same 80 OA subjects, grouping according to CPT<15°C (Table A4.3, Appendix 4). This analysis showed that those with lower mean CPT (<15°C) also tended to have an increased effect size and Cohen’s d value compared with those with high CPT (>15°C).
## Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Accepted/Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subjects taking etoricoxb will demonstrate:</td>
<td></td>
</tr>
<tr>
<td>a. Significant improvements in pain, dysfunction and quality of life.</td>
<td>Accepted</td>
</tr>
<tr>
<td>b. A significant reduction in measures of local hyperalgesia</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>c. A significant reduction in measures of widespread hyperalgesia</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>d. Significant improvements in self-report neuropathic pain score</td>
<td>Accepted</td>
</tr>
<tr>
<td>2. The measures of cold pain threshold and ADI will show</td>
<td>Rejected</td>
</tr>
<tr>
<td>similar responsiveness to the intervention of etoricoxb</td>
<td></td>
</tr>
<tr>
<td>60mg for 14 days.</td>
<td></td>
</tr>
<tr>
<td>3. Subjects with cold hyperalgesia (ADI ≥5) will respond less well to</td>
<td>Partially accepted</td>
</tr>
<tr>
<td>etoricoxb than those with a less hyperalgesic response (ADI &lt; 5).</td>
<td></td>
</tr>
</tbody>
</table>
9.6 Discussion

The aim of this study was to evaluate whether a standard Cox-2 anti-inflammatory intervention influenced measures of centrally-augmented pain in individuals with knee OA, or whether its effect was limited to local pain measures. A sub-analysis also evaluated whether those with more established centrally-augmented pain responded less favourably to the intervention.

**Improvements in pain, function and quality of life**

As reported in previous studies, etoricoxib significantly improved pain, function and quality of life in this knee OA cohort. Mean VAS pain for the last seven days reduced by 22% for those on active intervention compared with a 7% increase for the placebo group. WOMAC score for function-specific pain reduced even more in the Active group (30.6%) (Figure 9.2). Self-reported function improved similarly, with Active scores improving by 28.4%. The change in self-reported stiffness was most marked, with a decrease of 44% for Active group participants but an increase of 9% by Placebo participants, suggesting a significant impact on local inflammatory activity. There are only a few previous studies using etoricoxib 60mg with osteoarthritis patients and each has applied slightly different analyses (OMERACT-OARSI responder rates, effect size, NNT). However, the efficacy of etoricoxib as shown in the current study is consistent with previous data. (Lin et al., 2010) reported that 52% of subjects with OA knee and hip who were prescribed 60mg etoricoxib once daily for four weeks, decreased their WOMAC pain scores during walking by ≥ 30% (absolute mean reduction in score of 28%). In the current study 55% of subjects reduced their WOMAC pain sub-score by 30% and mean reduction in score by Day 14 was 30.6%. Pelosi et al. (2009) performed a pooled analysis of seven trials in OA hip or knee and reported an NNT of 3.5 for WOMAC pain reduction of ≥ 30% with etoricoxib 60mg. For the current study, the same analysis resulted in an NNT value of 2.9 (22/40 Active responders, 8/40 Placebo responders). Lin et al. (2010) also reported reductions in WOMAC-stiffness (23.1%) and function (25%) sub-scores. The current study found considerably greater reductions in stiffness but similar functional improvements. Improvements in SF36 quality of life values for physical and mental sub-scores found in the current study were also very similar to that reported by Lin et al (2010), with both studies finding improvements of 13 to 14% in physical functioning (PCS) and 7% in mental health (MCS).
The current study also tested the effect of etoricoxib on physical function, finding significant improvements in time taken to complete each task by the Active group whereas Placebo participants slowed slightly by Day 14 (Table 9.5). The stairs task differentiated most effectively between treatment groups, with Active participants improving their time by 13.2%, although this was still six seconds slower than the healthy controls of Study 5 (Figure 7.11, Chapter 7). Each of the three physical tasks cumulatively increased pain for all OA subjects, both at baseline and also at Day 14 (Figure 9.4a-b). Pain during the stairs task was again the most discriminatory, with Placebo participants reporting almost twice that at baseline whereas Active participants reported an improvement of 21%. Harden et al. (2009) reported that pain increase for a stairs task together with widespread PPT and thermal wind-up were key correlates of central sensitisation in OA knee subjects. On further analysis, the current study also found a very modest association between both local and widespread PPT and VAS pain for stairs task at Day 14 (PPT: mean r = -.322, p = .004; OA knee r = -.343, p = .002). These are only very weak correlations in a study not designed to test the hypothesis of Harden and colleagues. A future study comparing pain changes during physical tasks for larger groups of those with and without signs of centrally-augmented pain would be interesting for further clarification. There are no other pharmaceutical intervention studies with which these results can be compared Indeed the OMERACT-OARSI group has strongly advised that future OA studies should include a physical function task alongside self-report measures (Hawker et al., 2008).

**Local changes in hyperalgesia**

It was hypothesised that those taking etoricoxib would show significant reductions in local measures of hyperalgesia since the mechanism of action of a Coxib is predominantly peripheral (Brooks & Kubler, 2006). Subjectively, those taking the Active intervention reported a marked 44% reduction in stiffness, suggesting a significant impact on local inflammatory activity. It has also been proposed that PQAS ‘deep’ sub-score, which includes words such as throbbing and aching, also reflects inflammatory activity (Victor et al., 2008). Figure 9.8 illustrates that the Active group reported a change in sub-score very similar to that for stiffness: 45% reduction in contrast to an increase of 28% in these symptoms by Placebo participants. Alongside these self-report measures, significant improvements were also seen in the QST measures of PPT, PcPT and CPT for the Active intervention group. Mechanical hyperalgesia reduced by 32% for the Active group so that by Day 14 their PPT value was similar to that for the healthy cohort in Study 5 (Chapter 7, Figure 7.9). Placebo
PPT remained 30% more sensitised than controls. Punctate pain threshold also significant improved, although the margin of improvement was only small. Cold pain threshold at the OA knee also improved (Figure 9.5). This amounted to an approximate 10% reduction in sensitivity for the Active group and a 12% deterioration for the Placebo group. Heat pain threshold showed no significant change at any site, however, baseline HPT values were similar to those recorded for normal controls in Chapter 7. As for the 40 OA subjects with whom the controls were compared, this larger cohort of OA subjects did not exhibit significant levels of heat hyperalgesia at baseline.

This study therefore provides evidence that a two-week course of etoricoxib reduced subjective stiffness and inflammatory-type pain as well as local mechanical, punctate and cold hyperalgesia. This study confirms that COX-2 inhibitors influence peripheral inflammatory-induced hyperalgesia in humans. The enzyme cyclo-oxygenase 2 (COX-2) is strongly upregulated during inflammation and triggers the first stage of the catabolic conversion of arachadonic acid to prostaglandin E2 (PGE2). Complex positive feedback loops via PGE2 ensure the ongoing up-regulation of COX-2 expression, suggesting a role for COX-2 in more chronic inflammatory states (Bingham et al., 2006). In a study of humans undergoing hip replacement surgery, Renner et al. (2010) demonstrated that administration of etoricoxib pre-operatively resulted in high levels of drug in blood and wound fluid plus reduced levels of prostaglandins and IL-6 by the end of surgery (2 hours). Patients randomised to etoricoxib also required less pain relief in the immediate post-operative period. Only one previous small study has assessed the impact of an anti-inflammatory on evoked stimuli in humans with OA. Parks et al. (2011) used a painful pressure stimulus during fMRI imaging with six knee OA participants to analyse the effect of a two-week course of valdecoxib. However, the study reported that although increased levels of drug in blood and CSF correlated with improvements in spontaneous pain, local pressure-evoked pain did not change. The small sample size may be sufficient to explain the difference in result to the current study. However, this fMRI study also underlines that evoked and spontaneous pain do not necessarily reflect the same mechanism (Parks et al., 2011) and that changes may not occur concurrently.

**Widespread changes**

The current study also hypothesised that a Cox-2 inhibitor might have some influence on measures of central sensitisation. Animal studies have shown that COX-2 is present in CSF and can be influenced by COX-2 inhibitor intervention (Bingham et al., 2006).
Human studies also have reported the presence of COX-2 inhibitors in the CSF, implying the ability to cross the blood-brain barrier and influence inflammatory mediators in the CNS. Renner et al. (2010) and Parks et al. (2011) have reported the presence of effective levels of COX-2 inhibitors in CSF. Renner and colleagues administered a 120mg intravenous bolus of etoricoxib pre-operatively to patients undergoing joint replacement surgery and found therapeutic levels of the drug in samples of exudate, plasma and CSF taken 2 hours post administration. In a second similar study, the same group reported significant reductions in post-operative analgesic requirements by those randomly given pre-operative etoricoxib (Renner et al., 2012).

In the current study, there were some signs of a centrally-mediated effect by etoricoxib but it was by no means clear-cut. In addition to local changes, mechanical hyperalgesia at the unaffected knee was reduced by 14% but there was no change at the elbow. Punctate hyperalgesia was also reduced at the unaffected knee in addition to the OA knee, although there was no change at the elbow. However, lack of widespread effect cannot be concluded because, for both PPT and PcPT, elbow, baseline values were similar to that for the healthy controls in Chapter 7 and so not hyperalgesic. The same was found for HPT values, which were at a "normal" level at each test site at baseline so change following intervention would not be anticipated. Mechanical and punctate hyperalgesia that spreads beyond the affected joint, along with additional signs of activity-dependent central sensitisation, have been shown to be reversible in previous studies. Moss et al. (2007) reported the immediate increase in PPT at the affected knee in individuals with OA and also at the unaffected knee and ipsilateral heel following a single nine-minute application of a physiotherapeutic accessory mobilisation technique. Static and no-contact control conditions decreased PPT. PPT at the OA knee improved by around 26%, whilst that at the unaffected knee improved 15% and the control heel site improved by around 10%. Both Kosek and Ordeberg (2000) and more recently Graven-Nielsen et al. (2012) have shown that mechanical hyperalgesia as well as DNIC/CPM and temporation summation improve significantly for most patients after total hip or knee replacement. It has been proposed that reduction in peripheral nociceptive barrage gradually reverses activity-dependent central sensitisation, thereby reducing augmentation of pain signals and enabling dysfunctional descending inhibitory systems to normalise (Woolf, 2011; Graven-Nielsen et al., 2012).

CPT for the Active group however, showed almost no change at either the unaffected knee or elbow. This was true for all subjects, whether they showed significant CPT-cold
hyperlgesia at baseline or whether their baseline CPT was within normal limits. Very few studies have evaluated the effects of any intervention on cold pain thresholds. Chieng and Chan (2003) reported that TENS was ineffective and Sterling et al. (2010) found that a physiotherapy cervical glide technique was similarly ineffective at changing elevated CPT in those with chronic WAD, although cervical PPT and nociceptive flexion reflect did significantly improve. Behavioural cold hyperalgesia in humans has proved to be extremely difficult to explain mechanistically. This is likely to reflect in part the challenges of reliable cold pain measurement, as discussed in Chapter 2: 2.5, resulting in few published studies. It is equally likely to reflect the current uncertainties about the exact mechanisms involved. An abnormally noxious response to a normally non-noxious cold temperature appears to involve the complex interaction between many peripheral factors relating to transduction and transmission, central factors relating to reciprocal balance between heat and cold signals and supra-spinal factors involving descending inhibition and cortical interpretation.

However, in this study cold response was also evaluated using measures of intensity and sensation quality. Closer examination showed significant changes in these measures and this may help elucidate some of the mechanisms involved in a cold hyperalgesic response. Menthol ADI score, composed of VAS intensity values and an index value reflecting choice of unpleasant words to describe the quality, showed significant change following etoricoxib treatment. Sub-analysis showed that reduction in intensity of unpleasantness by 45% and pain by 60% were important drivers of change in total score for those in the Active treatment group, together with a 10% reduction in choice of the high-scoring word burning. VAS intensity and quality was also measured at CPT for each subject at each site. When mean combined values for all sites were reviewed, descriptor score (MWS) at CPT also showed significant change between Day 0 and Day 14 (t(79) = -1.94, p = .050). Although less significant than for ADI, it is significant to note that reduction in score for those in the Active group was once again driven by decreased choice of the word burning, by a strikingly similar 11%. There was no change in the intensity of cold experienced during either menthol or CPT testing.

So, there is a hint of a suggestion that, although CPT temperature did not change, there was the beginnings of a reduction in the noxious quality of the sensation at that temperature and this correlates clearly with the results found for the menthol test. It is interesting to speculate on the mechanisms that may be involved. Unpleasantness or
pain and a burning sensation during a cold stimulus imply activation of nociceptors. A reduction in intensity of these sensations in the presence of reduced peripheral and central sensitisation, as evidenced by changes in QST and self-reported measures as previously discussed, may be an additional reflection of reduced signal augmentation at the dorsal horn as a result of anti-inflammatory-induced reductions in central sensitisation. Reductions in burning sensations may also reflect de-sensitised heterosynaptic potentiation, although this might be expected to also cause a concomittant increase in cold sensation, which was not found. Alternatively, reduced unpleasant burning may reflect the reduced involvement of TRPA1 as circulating inflammatory mediators decline. TRPA1-expressing c-fibres have been shown to signal unpleasant burning sensations, associated possibly with their co-expression with TRPV1, but do not signal cold (Knowlton et al., 2013). TRPA1 has also been shown to mediate mechano-sensitivity when sensitised by a range of irritants, including inflammatory mediators (Bautista et al., 2012). The concurrent reduction in PPT and PcPPT supports the hypothesis that reduced TRPA1 activation may play a role in the less noxious quality of the cold sensation after anti-inflammatory treatment.

**Changes in neuropathic-type pain**

Self-reported neuropathic pain-type symptoms also showed significant improvements following anti-inflammatory treatment. PainDETECT score decreased by 23%, a magnitude similar to that for ADL. This was mirrored by an equivalent 29% reduction in PQAS ‘surface’ sub-score, which records qualities most closely associated with neuropathic-type pain. Change in PainDETECT neuropathic category was also significant for the Active group, with nearly 30% reducing their score sufficiently to move from ‘neuropathic’ or ‘unclear neuropathic’ to ‘negative neuropathic’ category. Few published studies have reported PainDETECT as an outcome following intervention and so there is little with which to compare these results. Gwilym et al. (2010) also reported a significant reduction in score following hip replacement surgery, although pre-operative scores were in the ‘negative neuropathic category’. Rados et al. (2013) report use of PainDETECT to evaluate effectiveness of lumbar epidural analgesia but no specific data is available for PainDETECT score change.

However, the current study findings provide additional data for the debate about whether positive self-report of neuropathic pain-type symptoms reflects undiagnosed neuropathy or central pain changes. A number of studies have reported self-reported neuropathic symptoms in people with hip and knee OA (Hochman et al., 2011;
Effect of etoricoxib in subjects with knee OA

Shigemura et al., 2011; Ohtori et al., 2012). A positive PainDETECT score may reflect otherwise undiagnosed neuropathy, caused potentially by articular denigration in OA. Alternatively, neuropathic pain quality to chronic pain may reflect a centrally-sensitised system in which there is widespread sensitisation and potentially some degree of neuroanatomical reorganisation. This could be due to changes in central biochemistry or descending signal inhibition, or altered central processing of neural information. The finding in the current study that self-reported neuropathic-type pain qualities can significantly reduce over a two-week period with an anti-inflammatory intervention demonstrates that for many individuals the symptoms are reversible. This is unlikely to indicate nerve damage for the majority. However, there were a few subjects in the Active group whose membership of the ‘positive neuropathic’ group was not influenced by etoricoxib intervention. It may be that this group of individuals have neuropathic-type pain that is driven by processes which may include less reversible central changes. An additional analysis was therefore carried out to evaluate whether the group of subjects in the ‘positive neuropathic’ category at Day 14 (8 who had not changed category plus 4 Placebo subjects whose score increased to ≥19 by Day 14) showed additional signs of centrally augmented pain. Table 9.14 below shows that those with ‘positive neuropathic’ scores showed significantly increased values for CPT, ADI and widespread (elbow) PPT. In particular, this ‘neuropathic positive’ group exhibited cold hyperalgesia as defined both by forearm ADI score ≥5 and CPT>15°C at all sites.

Table 9.14: Difference in CPT, ADI and elbow PPT values between those categorised as ‘positive’ or ‘negative’ neuropathic for PainDETECT at Day 14.

<table>
<thead>
<tr>
<th></th>
<th>+ve neuropathic</th>
<th>-ve neuropathic</th>
<th>t(78)=</th>
<th>p=</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(≥19) n=12</td>
<td>(&lt;13) n=51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPT OA knee</td>
<td>21.7 (.99)</td>
<td>9.5 (1.3)</td>
<td>-4.56</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Unaff knee</td>
<td>18.7 (1.8)</td>
<td>8.2 (1.1)</td>
<td>-4.28</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Elbow</td>
<td>21.4 (1.2)</td>
<td>8.9 (1.3)</td>
<td>-4.69</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Forearm ADI</td>
<td>5.4 (41)</td>
<td>3.2 (23)</td>
<td>-4.28</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>PPT elbow</td>
<td>275.4 (34.5)</td>
<td>380.7 (18.2)</td>
<td>2.56</td>
<td>.013</td>
</tr>
</tbody>
</table>

An additional factor of note is the significant increase in neuropathic-type symptoms by the Placebo group by Day 14. For example, PainDETECT score increased by 14% and PQAS ‘surface’ sub-score increased by a notable 68%. Such as large increase in neuropathic-type pain qualities is hard to explain despite a similar, although less marked worsening for AFL and CPT values for the Placebo group. Imaging studies have
shown that expectations play a significant role in influencing descending control of inhibitory or facilitatory pain mechanisms (Tracey, 2010) and it must be acknowledged that expectation may well have been a confounding factor for both Placebo and Active participants. However, subjects were blinded to treatment group and although they may have developed expectations of effect based on their personal judgement of group allocation, correct and incorrect expectations should have balanced the final result.

**Comparative response of ADI and CPT**

It was hypothesised that ADI and CPT would show similar responsiveness to the study intervention. However, this was not found to be the case and whilst ADI showed similar responsiveness to the intervention as PainDETECT, CPT temperature showed very little change. Although NNT for 30% improvement in mean CPT was only slightly less good than for ADI and PainDETECT, effect size and percentage change were considerably less (Table 9.9). It is of course possible that ADI and CPT are measuring different aspects of cold pain, but it may also be a reflection of the difference in effect size between QST measures and measures that involve self-report of quality. Effect size and NNT for PPT was a similar magnitude as for CPT. Imaging studies have reported that spontaneous pain (as measured by PainDETECT) is coded quite differently in the brain to pain evoked during CPT and PPT (Parks et al., 2011). It may be that, in measuring an evoked pain with the use of spontaneous pain-type words, the ADI bridges this apparent gap. There is little evidence to support or refute this idea. Gwilym strongly associates cold pain response with thalamic activity (2010) and other studies report strong activation of thalamic and limbic structures during spontaneous pain (Baliki et al., 2008). However, a future study using imaging would be needed to clarify whether ADI brain response more closely matches that of PainDETECT or of CPT.

**Heterogeneity of response to etoricoxib intervention**

Use of measures where severity groupings can be identified, such as PainDETECT, ADI or CPT, underlines that OA is not a homogeneous disorder. Study 6 (Chapter 8) showed that those identified as cold hyperalgesic with ADI score ≥5 also had a cluster of the most impaired scores for function, pain, widespread hyperalgesia and report of neuropathic-type pain. Based on previous studies suggesting an association between cold hyperalgesia and poor response to treatment, the current study had hypothesised that those with ADI-cold hyperalgesia would respond less well to etoricoxib. This was rationalised on the basis that chronically altered pain processes would not respond to a predominantly peripherally-acting anti-inflammatory.
However, the results of this sub-analysis were not as clear-cut as anticipated. Those with ADI-cold hyperalgesia responded positively to the majority of measures although to a lesser extent than the low ADI group. Self-report pain and function results were similar, with Active participants in both high and low ADI groups showing improvements in pain of around 30%. However, reductions in self-reported symptoms of inflammation were greater in the low ADI group, suggesting slightly greater efficacy in anti-inflammatory effects: WOMAC stiffness and function both improved approximately 15% more for the low ADI group and PQAS ‘deep’ score reduced by 20% more. QST measures were mixed. Although PPT change was similar between ADI groups at the OA knee, ADI-cold hyperalgesic subjects showed little change in widespread PPT, in contrast to the clear pattern of widespread improvement in the low ADI group (Figure 9.15a-b). CPT temperature showed no significant changes at any site for either of the ADI groups.

PainDETECT and ADI scores showed a reverse pattern to all other measures with, for both measures, the high ADI group improving slightly more than the low ADI group. For PainDETECT, 7/25 subjects (28%) with high ADI reduced their score sufficiently to be categorised as non-neuropathic by Day 14, in contrast to only 1/55 in the low ADI group. This was less marked for ADI scores although cold hyperalgesic subjects reduced their score slightly more than the non-cold hyperalgesic group: 11% versus 6%. Effect size and Cohen’s d mirrored these findings (Table 9.13). Table A4.3 in Appendix 4 shows the results of a similar sub-analysis undertaken using mean CPT<15°C as the cut-off value. Although this resulted in a considerably larger cold hyperalgesic sub-group, the results support the ADI findings.

It is unwise to draw any definitive conclusions from this sub-analysis of cold hyperalgesic versus non-hyperalgesic individuals. Since the study was powered for equal Placebo-Active groups and not for a further sub-grouping, power was low in the sub-analyses, and so Type I errors are likely. However, there are hints in this data that individuals with high levels of cold hyperalgesia may respond less well to standard anti-inflammatory interventions. Given the clustering of cold hyperalgesia with less good function, pain and quality of life and greater incidence of widespread hyperalgesia and neuropathic-type symptoms, it might be reasonable to propose that more central mechanisms are driving pain for these individuals. However, anti-inflammatory interventions do play a role in reducing nociceptive barrage and subsequent central
sensitisation, as many cold hyperalgesic individuals on Active medication improved in all measures to some extent. Therefore amongst those with cold hyperalgesia is a group whose symptoms are still malleable. There was also a group who did not respond to the intervention and whose unchanging positive neuropathic pain scores were associated with high CPT, ADI and widespread mechanical hyperalgesia. This may be the group that has cortical or brain-stem and spinal structural changes in pain processing, more akin to those who have well-entrenched neuropathic disorders. For these individuals, merely changing the peripheral or central sensitivity state is insufficient to change their pain because it is perpetuated by structural modifications in the CNS.

However, this study only provides initial data that hints at this hypothesis. Future studies, with more appropriate power, are needed to compare those with and without cold hyperalgesia. Ideally such a study would combine menthol-ADI testing with imaging and traditional QST, PainDETECT and functional measures, to evaluate whether a centrally-augmented group can be identified from those with more reversible centrally-sensitised pain. If a strong association with cold hyperalgesia and centrally-augmented pain is found, there are clear implications for clinical assessment of cold hyperalgesia and evaluation of more appropriate treatments for these individuals. In addition, there are strong implications even from the current study that for future OA studies need to acknowledge heterogeneity and modify inclusion criteria accordingly. A study that were to evaluate the effects of a centrally-acting neuropathic pain intervention on individuals with OA showing signs of centrally-augmented pain, would also assist in clarifying the mechanisms involved and potentially provide a better treatment option for those with such intractable pain.
9.7 Summary

In summary, this double-blind, randomised, placebo-controlled study investigated the impact of a 14-day course of etoricoxib on local sensitivity and centrally-mediated pain in knee OA subjects. It was also hypothesised that those OA subjects identified in Study 6 (Chapter 8) as cold hyperalgesic (ADI ≥5) would not respond as well to etoricoxib as those with more normal cold responses. Sub-analyses were therefore run, comparing Active and Placebo responses within ADI groups.

The main study found that etoricoxib was significantly more effective than placebo intervention in reducing self-reported pain, function and quality of life. Etoricoxib also significantly reduced local mechanical and cold hyperalgesia. Additional large improvements in self-reported WOMAC stiffness and in PQAS ‘deep’ sub-scale supported the likelihood that etoricoxib had a positive impact on local pain sensitivity. Although heat and punctate pain thresholds showed no change, baseline values for both were already within normal limits.

In addition, there was evidence that etoricoxib influenced measures of centrally driven pain. There were clear improvements for the Active group in PQAS paroxysmal and surface scores. For PainDETECT, by Day 14, ¾ of Active group subjects had reduced their score sufficiently to be categorised as ‘negative neuropathic’, having previously been classified as either ‘unclear’ or ‘positive’. There were significant reductions in mechanical and punctate hyperalgesia at the unaffected knee, although not at the elbow. Widespread cold response, as measured by forearm ADI score, significantly improved in the Active group, reducing unpleasantness and pain intensity ratings and less frequently selecting the words burning and prickling. In contrast, CPT did not change at either the unaffected knee or the elbow.

It was hypothesised that there may be a differential response to a Cox-2 anti-inflammatory intervention, based on cold hyperalgesia. Active-Placebo differences were therefore re-analysed separately for each ADI group. Although Active participants in both ADI groups showed significant improvements in outcomes in line with the whole group analysis, those with low ADI tended to show slightly greater improvements than those with high ADI scores. In particular measures of local inflammatory effect, such as ‘deep’ PQAS sub-score and WOMAC stiffness score, were more improved in the low ADI group. The exception was for ADI and PainDETECT scores, which improved slightly more in the high ADI group. Analysis of effect size
reflected a similar pattern, as did an equivalent analysis using a CPT<>15°C division. However, these results can only be considered an indication of a possible pattern since power was and the high ADI group was relatively small.

This study provides evidence for the ability of a Cox-2 inhibitor to influence both local and centrally-mediated pain signs in individuals with OA. However, response is not homogeneous and the presence of cold hyperalgesia may influence how strong that response is. ADI values were more responsive than CPT temperature in this study and are therefore suggested as a valid measurement tool to differentiate cold hyperalgesic from non-hyperalgesic individuals in future studies.
Chapter 10

Discussion

The overall purpose of this investigation was to develop a clinically-relevant test for cold hyperalgesia that could be used to identify individuals with features of centrally-augmented pain. The investigation showed that the topical menthol test with associated ADI scoring system was a reliable method for identifying cold hyperalgesia in both healthy and chronic pain cohorts, that showed good content, construct and criterion validity.

Dose-dependent effects of menthol

The first two studies evaluated whether graded concentrations of menthol evoked graded sensory responses in healthy subjects. These studies showed conclusively that the sensory system was able to differentiate between menthol concentrations both in terms of sensation intensity and sensation quality. Higher concentrations of menthol evoked higher intensities of cold, heat, unpleasantness and pain as well as higher pain descriptor scores. This was found consistently across both liquid and gel formulations, to the extent that the words chosen at equivalent concentrations were remarkably similar. A normal response at low or high menthol concentrations could therefore be characterised quite clearly. A low concentration was more likely to be described as cool or tingling and rarely unpleasant, although sometimes accompanied by low intensity warmth, whereas higher concentrations tended to evoke unpleasantness and possibly pain (although at a low levels of intensity) and more intense descriptors such as icy/freezing, burning and stinging.

Increased intensity and more noxious quality of response at higher menthol concentrations strongly mirrors the effect that colder temperatures have on the peripheral sensory system, suggesting that similar mechanisms are activated by the different modes of stimulus. Davis (1998) found that low, briefly sustained cold temperatures (2°C) evoked increasing report of unpleasantness or pain accompanied by burning and intense dysaesthetic qualities such as prickling. Chapter 5 of the current investigation found similar sensory responses in healthy subjects to a less cold temperature (10°C) sustained for a slightly longer time-period. Comparison between response to sustained 10°C (Chapter 5) and response to menthol gel B (Chapter 4)
using similar intensity and quality measures in two separate cohorts of healthy subjects found remarkable similarities, particularly in terms of the choice of descriptors.

These psychophysical findings correspond with basic science investigations into the mechanisms by which increasingly noxious (low) cold temperature is signalled. As cold temperature lowers, increased influx of calcium into TRPM8-expressing Aδ thermal fibres begins to slow and calcium influx into low-threshold TRPM8-expressing Aδ- and c- nociceptive fibres increases (Madrid et al., 2009; Latorre et al., 2011). This barrage of cold signals at the dorsal horn is likely to be interpreted as increased intensity. However, as cold temperature drops further the sensation quality begins to change and unpleasantness is felt in addition to cold. It has been proposed that this is an indication that higher threshold Aδ nociceptors have increased their activation and that the sensations of mildly noxious stinging or burning may reflect a transitory warning signal (Campero et al., 2009). However, cold sensation continues to be experienced for most individuals. Knowlton et al. (2013) have recently demonstrated that the sensation of cold is transduced almost entirely by TRPM8 channels so this implies that TRPM8 channels are still responsible, even at low cold temperatures, although it is still unclear whether the cold sensation is signalled via TRPM8-expressing nociceptors (McCoy et al., 2011). No similar work examining concurrently the basic science data and equivalent psychophysical experience appears to have been completed using graded concentrations of menthol. However, the similarities in sensory characteristics between increasing intense cold and higher menthol concentrations in the current investigation strongly suggest a similar neurophysiological process.

Despite the similarities in overall effect, temperature-cold and menthol-cold stimuli clearly differed in response pattern. Cold temperature reached peak intensity within in the first minute (Figure 5.9a) and then steadily declined within the five-minute application. In contrast, response to menthol was gradual, with several minutes lag time before most individuals reported any sensation. Peak sensation was reached after nine to 12 minutes, depending on sensation and then just started to decline before removal of the stimulus at 15 minutes (Figure 5.9b).

Although the basic science literature does not provide an explicit mechanism for this timing difference, variations in permeation, diffusion and channel binding mechanisms are likely to be the cause. The lipophilic nature of menthol enables it to permeate the skin with great efficiency so that the chemical is often added to topical preparations to
enhance permeability (Fang et al., 2008). However, the diffusion capacity of cold
temperature is considerably greater, enabling cold to reach the epidermis with almost
no lag time. In the presence of the molecule PIP₂, TRPM8 channel voltage change occurs
directly in response to cold temperature, resulting in immediate calcium influx with no
secondary messenger involvement (Knowlton, 2011; Yudin & Rohacs, 2012). It is an
immediate, ‘all or nothing’ response so that cold sensation is felt immediately. In
contrast, menthol activation of TRPM8 occurs gradually and requires the involvement
of intra-cellular cascades. TRPM8 has four binding sites for menthol. As each site is
filled, internal cascades are initiated, calcium influx gradually increases and the channel
open state becomes more stable (Janssens & Voets, 2011). As more menthol permeates
to the extra-cellular environment the increase in number of activated TRPM8 channels
would correspond with the gradual build up in intensity of cold sensation that menthol
evokes. Interestingly, this distinctly different onset pattern for temperature and
menthol is also seen for heat, unpleasantness and pain. This may mean that the key
factor in timing pattern is permeation speed through the epidermis. Equally it may be a
sign that the mildly noxious sensations that gradually intensify alongside cold are from
predominant activation of lower threshold TRPM8-expressing nociceptors rather than
from higher threshold c-fibres. The slow steady decline of cold temperature following
the initial peak is likely to reflect the habituation mechanisms described during in vitro
studies. Once calcium influx reaches a critical level, the PIP₂ molecule is dislodged via
second messengers and cold-induced TRPM8 voltage change is disabled. Menthol
deactivation mechanisms have not been so clearly but are likely to exist.

Although a brief cold stimulus is conventionally used to assess cold response, use of a
sustained cold temperature has advantages. The individual being assessed has greater
time to consider the accuracy of their response and a sustained stimulus provides an
additional temporal variable. Although sustained cold delivered by a thermode (as used
in Chapter 5) has the advantage of enabling a variety of temperatures to be set, it has
the disadvantage of equipment cost. Application of ice has been proposed as an
inexpensive way to provide a sustained cold stimulus (Maxwell & Sterling, 2012).
However this modality is problematic because the intensity of noxious input is difficult
to control, cannot be modulated and may evoke additional vascular nociceptive input,
thereby intensifying the sensation still further. Topical menthol in contrast provides a
non-thermal sustained cold stimulus of infinitely variable intensity according to
concentration.
The findings from these first studies suggest that the overall effect of menthol is remarkably similar to that of cold temperature when equivalent concentrations and temperatures are compared. This similarity is seen in terms of sensation intensity and sensory quality. This strongly suggests that a sustained menthol stimulus may be a valid sensory equivalent to a sustained thermal stimulus.

**Sensitivity and validity of the ADI scoring system**

Studies 1 and 2 demonstrated a response measurement approach that included both intensity and quality domains provided better content validity. Although the McGill PRI score offered a useful initial descriptor measurement approach, its rankings were designed to characterise chronic pain symptoms and so it was not ideal for identifying evoked cold hyperalgesia. The ADI scoring system was therefore developed to quantify response to the menthol-cold stimulus. The sensory domains of intensity and quality were combined in the score in order to provide more comprehensive characterisation of cold response and a weighting system developed, based on findings from Studies 1 and 2 as well as previous studies in healthy and OA cohorts (Moss et al., 2010; Wright et al., 2010a). Chapter 5 evaluated the ability of the ADI scoring system to discriminate between thermal-cold temperature using a contact thermode. This study found that the total ADI was able to discriminate clearly between temperatures, a lower temperature resulting in a higher score. Each component of the ADI individually discriminated between temperatures, indicating that the elements of intensity and quality were equally valuable in characterising response. In addition, analysis of each component enabled analysis of the mechanisms by which cold response was mediated. The anticipated increase in VAS intensity with decreased temperature was likely to reflect increased activation of cold-sensitive neurons via voltage-dependent calcium influx and opening of potassium channels such as Nav1.8. All intensity ratings correlated well between temperatures, indicating internal consistency of both the measure and the stimuli.

The VAS ratings used for the ADI were not just for pain, which is often the only intensity measured by other studies. Additional VAS data for cold, heat and unpleasantness were included to reflect the reported effects of noxious cold (Davis & Pope, 2002). The broader range of intensity data also enabled a far better picture of a complex sensory response to emerge. For example, increase in cold intensity without

---

1 Ref Appendix 5
any additional sensations might suggest activation of cold-thermal fibres. Increasing
cold accompanied by unpleasantness might indicate triggering of low threshold cold-
sensitive nociceptors, which may be Aδ or c2-fibres. These fibres express fewer TRPM8
channels but more Nav1.8, TRPV1 and TRPA1. Intense pain is more likely to reflect
activation of higher threshold c-fibres, which might then release additional pain-
activating peptides such as substance P. Sensation quality data provides additional
indications of the possible mechanisms activated by the cold stimulus. For example,
tingling selected alongside low intensity unpleasantness reinforces the likelihood that
low-threshold Aδ-fibres may be active, whereas choice of icy, burning or more intense
stinging highlights a situation in which c-fibres may be more dominant. Fruhstorfer
(1984) proposed that these particular words reflect specific activation of c-fibres,
although other studies have shown that they can be selected during Aδ-fibre activation
(Campero & Bostock, 2010).

There were suggestions in the studies that, by measuring both increased intensity and
change in quality, the ADI score was able to indicate more sensitively than CPT the
point at which an individual response becomes abnormal. Chapter 5 compared the
reliability and variability of the menthol ADI with that reported for CPT and found that
the ADI exhibited similar reliability but considerably lower variability in total score and
MWS descriptor sub-score. It may be that the less defined sensation of cold pain
requires a more multi-dimensional measurement system than CPT can provide. Cold
sensation and pain involves the complex interaction between membrane channels,
intra-cellular cascades and inhibitory-disinhibitory interactions between nerve fibres.
Cold pain is therefore less defined than heat pain, meaning that accurate identification
of a temperature at which the cold becomes painful is likely to be variable.
Theoretically, the menthol test allows for greater accuracy by providing more time to
evaluate response, and by providing measures that encompass both intensity and
sensation quality.

More comprehensive characterisation of normal response to increasingly unpleasant
temperatures or menthol concentrations also provides a mechanisms-based rationale
for abnormally elevated responses at milder temperatures. For example, for individuals
with OA, higher than normal intensity during an evoked non-noxious cold stimulus
(temperature or menthol) may indicate central augmentation, either from central
sensitisation or from inhibitory system dysfunction. Changes in quality from cold-type
sensations through tingling to burning and stinging or prickling may indicate the
progression from Aβ to c-fibre activation. An icy, burning sensation in response to a normally low intensity cold stimulus may indicate that the normal spinal inhibitory processes have been disrupted (Susser et al., 1998). Alternatively, if the paradoxical sensation is particularly marked or intense, this may indicate functional or structural phenotypic change in the dorsal horn (Ueda, 2006).

In the current investigation, the ADI scoring system demonstrated improved content validity over alternative scores such as the McGill Pain Rating Index (PRI). Studies 1 and 2 demonstrated that the sensitivity of a descriptor measure such as the PRI was enhanced with additional information about the intensity of particular sensory elements. Comparisons between the PRI and ADI in Chapter 5 showed that, although the PRI was unexpectedly sensitive to predicting membership of the cold hyperalgesic (CPT>15°C) group, the ADI showed equal sensitivity (.90) but improved specificity (.74). This excellent level of predictive ability for the ADI illustrated that, in addition to providing more comprehensive information, the scoring system was appropriately weighted to be able to identify cold hyperalgesia as measured by conventional CPT.

**Discriminative effects of topical menthol**
The combination of a stimulus that evoked differentiated effects dependent on concentration and a measurement system sensitive to change in either intensity or quality, meant that the menthol ADI test was able to differentiate quite clearly between normal and abnormal responses to a given concentration. For Gel B, the majority of subjects reported a predominantly cool or cold sensation, with some degree of low intensity unpleasantness. An atypical response was therefore clearly identifiable, involving higher intensity unpleasantness or pain and selection of the noxious words burning, stinging or prickling. Dichotomous cluster analysis for VAS and descriptor scores in Chapter 4 showed that those in the high clusters exhibited distinctly different responses to those in the low clusters with no group overlap (Table 4.4). This noxious response was also associated with higher CPT value and earlier sensation onset (Chapter 3), adding an additional temporal variable to the discriminative criteria.

In Studies 5 to 7 where larger cohorts of subjects with chronic pain were evaluated (n=4 and n=80), the menthol test showed discriminative ability equal to that of CPT. In Study 5 (Chapter 7), cold response, whether measured by menthol ADI score or CPT temperature, differentiated the OA group from the matched control group more effectively than other QST or functional measures. When total ADI score was divided
into sub-components, OA subjects showed significantly higher VAS intensity for cold and unpleasantness, although not for heat or pain (Figures 7.2a and b) and more frequently described the sensation as burning, stinging or prickling than healthy controls (Figure 7.5). Elevated widespread cold response therefore differentiated those with OA pain and although both CPT and the menthol test were able to identify hyperalgesic individuals, the ADI components provided additional information about the degree of sensitisation.

However, it was clear from Chapter 7 that not all OA subjects were cold hyperalgesic, so Chapter 8 investigated the extent of menthol-cold hyperalgesia in a larger cohort of 80 OA participants. Thirty-one percent of OA subjects scored ADI≥5, which corresponded relatively well with the 37% classified as cold hyperalgesic using the CPT cut-off of 15°C. These subjects also reported the highest levels of spontaneous pain and disability. This finding is supported by other studies using CPT to investigate cold hyperalgesia in OA (Moss et al., 2008; Wright et al., 2010b)1. Two additional studies did not support the presence of cold hyperalgesia in OA participants, although both studies were associated with the methodological issues of CPT testing discussed in Chapter 2: 2.5

CPT and the ADI identified cold hyperalgesia in both healthy (Studies 1 and 2) and OA cohorts (Studies 5, 6 and 7). For healthy subjects with no current pain and no history of pain condition, a noxious response to a stimulus that is considered non-noxious by the majority raises interesting questions about mechanisms. Healthy subjects with an atypical response reported similar types of sensory changes to those with OA: increased VAS intensities and altered sensation quality that included burning and stinging. In both healthy and OA subjects, the pain system had augmented the normal response, although at a lower level of intensity in the healthy group. In the case of pain-free subjects this cannot be explained by sensitisation secondary to injury or inflammation. For these subjects, widespread augmented and altered sensory response is more likely to reflect central changes in inhibitory or interpretative systems. Previous studies have shown that a proportion of healthy individuals exhibit inefficient descending pain inhibition (Granot et al., 2008), which might explain the increase in intensity. The changes to sensory quality may also be a reflection of altered inhibitory pathways but as a result of supra-spinally-mediated descending facilitation. It has been suggested that this modulation of PAG function may be genetic in origin or may result from past influences over prefrontal activity, as has been proposed for fibromyalgia (Clauw, 2009). Yarnitsky et al. (2008) have shown that pre-existing DNIC dysfunction
in pain-free patients is associated with amplified pain after surgery and proposed that genetics may be partially responsible.

It is likely also that some of those with OA may have had a dysfunctional endogenous pain modulation system that pre-dated development of their OA symptoms. This reduced inhibitory drive may have served to augment pain response to peripheral nociceptive input. It has been shown that for some individuals with OA, pain may not just be driven by peripheral nociceptive barrage but that dysfunctional central inhibitory processes may be driving high levels of spontaneous pain and elevated response to evoked pain (Arendt-Nielsen et al., 2010). Several studies have now reported that for some with OA, reduction in pain following total joint replacement is associated with return to normal levels of descending inhibition (Kosek & Ordeberg, 2000; Graven-Nielsen et al., 2012). This suggests that for these individuals tonic nociceptive barrage from their damaged knee influenced reduced inhibitory control rather than vice-versa. For about 20% of individuals, replacing the joint does not change the pain, implying that nociceptive barrage is not the driver. In these individuals, a centrally-driven process perpetuates pain and it may be that the PAG is actively facilitating spinal pain signals. However this is purely speculative and a future study is needed to test this hypothesis.

**The menthol test as an indicator of central-augmentation**

Centrally-augmented pain is therefore a more challenging condition to treat because by the time the symptoms have been identified, pain processes are firmly established and may involve structural changes. Earlier identification of individuals who are starting to develop signs of centrally-augmented pain would enable treatment to be redirected at an earlier stage towards more centrally-acting interventions. It has been hypothesised that cold hyperalgesia may be a particularly useful indicator of centrally-augmented pain. In contrast to other QST measures such as PPT, cold response tends to be consistent between body sites, suggesting a mechanism that is supra-spinally mediated. Fibromyalgia studies have consistently shown that whilst mechanical and heat hyperalgesia may only be shown at painful sites, cold hyperalgesia tends to involve non-painful sites also (Kosek et al., 1996; Desmeules et al., 2003). Cold hyperalgesia is also reported to be associated with higher levels of pain and disability and poor response to treatment in whiplash (Sterling et al., 2006; Jull et al., 2007) and in fibromyalgia (Hurtig et al., 2001).
Chapter 8 applied the menthol test and ADI scoring system to explore correlates of cold hyperalgesia in individuals with knee OA. The study found that those with high forearm menthol ADI score also exhibited widespread mechanical and heat hyperalgesia and high levels of pain and disability. The widespread nature of these indicators suggests a centrally-mediated mechanism, although activity-dependent central sensitisation rather than supra-spinal mechanisms may be the driver. So for some subjects, chronic peripheral nociceptive input from the damaged knee may stimulate changes in central pain processing in the dorsal horn, leading to central sensitisation possibly accompanied by activity-induced dysfunction in spinal inhibitory control. In this cohort mechanical, heat and punctate hyperalgesia were most elevated at the OA knee, implying that ongoing peripheral nociceptive input was at least partly responsible for pain augmentation in some individuals.

However, many of those with high forearm ADI also exhibited CPT>15°C at all three test sites and high PainDETECT score, both of which may be indicators of more established central pain augmentation (Gwilym et al., 2010). There is great debate over the meaning of a positive neuropathic score in an individual without a diagnosis of neuropathy. Thirty-two percent of those with high ADI in this study scored as positive neuropathic on PainDETECT, in contrast to only five percent of those in the low ADI group. High PainDETECT score in the current study was associated with greater choice of neuropathic-type descriptors in the Pain Quality Assessment Score (PQAS) and with more frequent report of burning and stinging sensations during menthol application. Whilst augmented or altered signalling may indicate damage to small diameter neurons, it seemed unlikely in this cohort who showed no signs of globally reduced nerve function. Indeed, other studies have reported that high PainDETECT score is associated with fewer anatomical or biochemical markers of damage (Ohtori et al., 2012). An alternative explanation for a high PainDETECT score associated with additional central pain markers is that the neuropathic-type symptoms are driven by more hard-wired changes at spinal and supra-spinal levels, which influence augmentation and interpretation of sensory signals. Gwilym (Gwilym et al., 2009; Gwilym et al., 2010) has reported thalamic grey matter changes and changes in periaqueductal grey matter activity that correlates with PainDETECT score in individuals with hip OA. This is suggestive of a link between neuropathic-type pain and changes to mid-brain inhibitory-facilitatory balance or interpretation.
However, the association of cold hyperalgesia to a centrally-augmented pain state is still speculative and difficult to prove or disprove conclusively using QST alone. The final RCT study in this investigation sought to clarify further whether the measures proposed as reflective of central augmentation were influenced in any specific way by a Cox-2 non-steroidal anti-inflammatory drug. There is some evidence that Cox-2 inhibitors may cross the blood-brain barrier (Renner et al., 2010) and act on central spinal terminals, but it is likely that their main action will be peripheral. So, although peripheral and central sensitisation may be reduced, it is unlikely that more hard-wired or structural changes in pain processing will be influenced by a short acting Cox-2 intervention.

The results from Chapter 9 were not entirely clear-cut, indicating an effect even within those with signs of centrally-augmented pain. Etoricoxib clearly improved measures of peripheral sensitisation such as local mechanical hyperalgesia, subjective stiffness and PQAS 'deep' subscale qualities. In addition, there was evidence that measures of more centrally-driven pain were also influenced. There were some signs of widespread improvements in mechanical and punctate hyperalgesia, although significant change at the elbow was hindered by the normality of baseline values for the Active group. Improvements in these measures corresponds with previous studies that have shown that both local and widespread mechanical hyperalgesia is reversible (Graven-Nielsen et al., 2012) even after a brief pain-relieving physical intervention (Moss et al., 2007). This may indicate influence of the anti-inflammatory medication over activity-dependent central sensitisation.

However, PainDETECT score, which has been proposed as a potential indicator of centrally-mediated pain also improved significantly in the Active intervention group. Twenty-five percent of those categorised as 'positive' or 'unclear' neuropathic at baseline reported sufficient reduction in score to be reclassified as 'negative neuropathic' by the end of the two-week study. Improvement in ADI score was similar, with those in the Active group showing significant reductions in unpleasantness intensity and reduced selection of noxious words such as burning by Day 14. However, it is important to note that, for both PainDETECT and ADI, not all subjects responded. Those whose ADI score at baseline was ≥5, still scored as cold hyperalgesic (≥5) at Day 14. Closer examination of PainDETECT also showed that those who were still 'positive neuropathic' at Day 14 also failed to improve in other key measures. At Day 14, this group continued to exhibit a mean hyperalgesic ADI score of 5.4, significantly elevated
CPT at every test site (mean 20.6°C, p<.001) and significantly sensitised PPT (Table 9.14). This indicates that, there was a sub-group of subjects with more marked signs of central pain augmentation, including high ADI score, whose symptoms were not reversed by etoricoxib.

In apparent contrast to ADI score, CPT temperature changed little at any site over the 14-day intervention. Effect sizes for CPT were also considerably lower than for ADI, with CPT mirroring the effect size for similar QST measures and ADI mirroring effect size for PainDETECT, which involves similar pain quality assessment. This difference could indicate that ADI and CPT measure different phenomena, although the basic science evidence does not support this. Alternatively the results may reflect once again differences in measurement sensitivity. The ADI scoring system allows for subtle changes in intensity or quality to be registered. Analysis of changes at Day 14 shows that sensations of cold and heat were not significantly altered in terms of intensity or words choice in the Active group. Changes in score were mediated mainly by reductions in unpleasantness and pain intensity (Table 9.5) and in selection of the words burning and prickling. This may indicate that a primary action of etoricoxib was to reduce activation of high threshold c-nociceptors, possibly mediated by reduced sensitisation of TRPA1 secondary to reductions in inflammatory mediators (Bautista et al., 2012).

CPT as a simple temperature value was unable to discern these subtle changes. Chapter 9 did however collect additional descriptor data for quality of sensation at CPT. The descriptor MWS score for CPT did indeed show some change by Day 14 in the Active compared with the Placebo group (p=.050). This change was driven largely by a decrease in selection of the word burning, by 11% of subjects in the Active group. Although of lesser magnitude, this change mirrors the 10% reduction in selection of burning sensation for the menthol test.

The ability of the ADI to identify subtle changes, or lack of change, in pain quality and intensity is a particularly valuable as it means that degrees of pain centralisation or augmentation can be identified. Individuals can then potentially be provided with the most appropriate intervention and efficacy monitored. For example, a small reduction in report of unpleasantness and burning for forearm ADI in an individual with knee OA following 14 days of a Cox-2 inhibitor would indicate that the drug was successful in reducing central pain mechanisms to some extent. If total ADI score was still >5
continuation of the Cox-2 treatment might well be indicated. An individual who reported no change from a high ADI score may have an augmentary drive that is supraspinally-mediated. Continuation of the Cox-2 is therefore not indicated, but a centrally-acting intervention might be more appropriate. An individual whose ADI score significantly improved over the two weeks to a non-hyperalgesic level (<5) may not need to continue with the Cox-2, perhaps being advised to use the drug in an intermittent targeted manner. This is clearly speculative but indicates the potential that an effective measure for widespread cold hyperalgesia might have in a clinical environment. There is clearly a need for future studies to examine whether this hypothesis is correct.

Study 7 (Chapter 9) had also hypothesised that those with baseline cold hyperalgesia would respond less well to the Cox-2 inhibitor intervention as a result of more entrenched and less reversible centrally-augmented pain. However, the results only hinted that this might be right. Those with cold hyperalgesia (defined as either CPT > 15°C or ADI ≥5) did respond to etoricoxib, but less well than non-hyperalgesic subjects in almost all parameters, once again suggesting that even amongst those with cold hyperalgesia there is a sub-group whose symptoms are still malleable. However, the study was not powered for this sub-analysis and Type 1 errors were quite possible.

In addition, and perhaps more importantly so far as future studies are concerned, within the cold hyperalgesic group there was a sub-group of Active subjects who at Day 14 still exhibited a cluster of elevated scores for ADI, PainDETECT, CPT and PPT. These subjects clearly did not respond to the Cox-2 intervention and so may be the group that had CNS structural changes, requiring more centrally-acting interventions. This finding raises several implications for the menthol test. The currently proposed ADI cut-off score of 5 may be sufficient to identify those with centrally augmented pain but may not be high enough to identify those who are resistant to treatment interventions. As with the PainDETECT scoring system, two levels of severity may be useful in differentiating degree of central effect on pain and also in monitoring extent of response to interventions. It may equally be that the menthol test on its own does not sufficiently characterise centrally-augmented signs. It has certainly been reported that evoked pain (such as the menthol test) and spontaneous pain provide different effects but brain imaging studies that combine both together offer a more comprehensive picture of functional and structural change (Parks et al., 2011). Combining the menthol test with a self-report tool that measures spontaneous pain intensity and quality, such
as PainDETECT or PQAS may therefore provide a more comprehensive and reliable measure of centrally-augmented pain.

**Limitations**

This investigation chose to focus on the initial evaluation of validity and reliability of the menthol test. Inevitably this has meant that additional factors that may have influenced the results have not yet been addressed. Now that basic validity and reliability have been demonstrated in the cohorts selected for this series of studies, future studies are need to evaluate the impact on response to menthol of additional factors such as psychological influences or cultural or linguistic factors.

Psychological factors may influence response to an evoked stimulus in individuals with chronic pain. A number of studies have found that diagnosis of depression changes pain perception, although it is unclear whether this results in increased or decreased response. Klauenberg et al. (2008) found that CPT and wind-up were both increased in pain-free subjects with depression and to a lesser extent in those with fibromyalgia but no depression. Pressure pain threshold and detection thresholds were not significantly different between groups. The authors concluded that depression was therefore associated with increased central excitability and so potentially may augment signs of hyperalgesia through dysfunctional inhibitory serotonergic mechanisms. Smart et al. (2012) also found that clinically-assessed signs of central sensitisation were associated with higher depression scores for individuals with chronic low back pain. Wylde et al. (2011) found that persistent severe pain following total knee arthroplasty was independently associated with diagnosis of major depression and this is supported by other studies into post-surgical persistent pain (Hinrichs-Rocker et al., 2009). Pain catastrophising has been associated with changes in central sensitisation measures. For example pain catastrophising has been associated with increased temporal summation in those with chronic low back pain (George et al., 2005) and with poor outcome after knee replacement surgery (Singh et al., 2011). It has been suggested that if serotonergic pathways are altered, pain is facilitated rather than inhibited (Klauenberg et al., 2008). Not all studies agree that psychological factors influence pain sensitivity. Park et al. (2010) found that those with arthrogenic TMJ disorder exhibited central sensitisation, as measured by widespread elevated QST scores, but this was not associated with higher somatisation and depression scores. The balance of evidence however suggests that psychological factors, in particular altered mental health and pain catastrophising, may be influential in modulating an individual’s response to
spontaneous pain or to evoked stimuli such as the menthol test, and therefore this needs to be investigated in more detail.

There are a range of follow-up investigations that are needed to ensure ideal clinical applicability for the menthol ADI test. The normative studies of this investigation showed some other limitations. Small cohorts of healthy subjects from a limited source were used for all studies and so the results of this investigation should be considered to be indicative rather than conclusive. Future studies in larger and more varied normal cohorts are now needed to clarify the characteristics of normal and abnormal response and so confirm (or otherwise) that the current ADI weightings are appropriate. For example, influence of age needs to be clarified. There were hints in the current studies that age may influence intensity rating. For example, Study 5 (Chapter 7) showed that older healthy adults rated intensity significantly lower than the healthy younger subjects of Study 2 (Chapter 4). Previous studies provide conflicting evidence regarding age. A recent study compared younger (mean 26 years) and older adults (mean 79 years) and reported a significant difference between groups in activation of particular brain areas during evoked pain and also in selection of descriptors for pain quality, although there was no difference in reported pain intensity (Cole et al., 2010) and this supported an earlier study by Gagliese and Melzack (2003). Yet Gagliese and Katz (2003) reported that post-operatively older subjects rated pain as less intense than younger counterparts. A study comparing intensity and quality of pain response to menthol in older and younger healthy subjects is therefore necessary to clarify this and confirm whether the current ADI weightings are appropriately scaled for all age groups.

Factors relating to culture and language clearly may also influence response to an evoked sensory test, particularly when the measurement tool involves selection of words. This study unintentionally applied the menthol test to a defined and limited cohort group both for the healthy and clinical participants: largely Caucasian, English-speaking and with a sufficient level of education to enable them to participate fully in a pain study. The MWS sub-score of the ADI replies depend on grasp of sufficient vocabulary in order to describe the sensory experience. The limited number of words chosen for weighting in the current ADI therefore may not necessarily be appropriate for different cultural or language groups. For example Indigenous Australians may not necessarily choose the same language to describe the sensations that the ADI seeks to identify. Even within one linguistic heritage, grasp and use of language may vary so that
certain pain words may not be universally used. On the other hand, key sensory words such as prickling, stinging and burning, which differentiate an abnormally augmented response may cross boundaries within English speaking cultures. This needs to be more fully investigated. Considerable work would also be needed to apply the test to different language groups, since certain words, such as tingling or prickling have no direct translation in other languages. In addition to language, pain response also involves complex cultural factors. Although the menthol test is designed to primarily require a sensory response, this will inevitably be influenced by additional cultural factors which would need investigation to ensure the most widespread applicability for the test.

Further reliability testing is also needed. The current investigation used only a small cohort in a simple two-occasion test-retest design. Although these results suggest that menthol response is a relatively stable measure, a study testing larger subject numbers over three or more test occasions would provide more conclusive results. In addition, the current test application time of 15 minutes is still impractical for the time constraints of some primary practitioners (for example, an average GP consultation in the UK is only 10 minutes). The current studies showed that peak response was reached by the majority of subjects within 10 minutes, so that the test could be shortened to improve practical application. However this needs to be clarified more specifically with larger cohorts.

Reliability and validity also need to be assessed at sites other than the volar forearm site used in the current studies. This site was used as it is generally not associated with pain pathology and, as a largely hairless zone for most individuals provides a more standardised site for transdermal delivery. However, further normative testing is needed to review whether response varies according to body site or skin type. The forearm site may not always be appropriate to use but, more importantly, the hypothesis that response to menthol will be similar at across body sites due to its proposed centrally-mediated mechanism, needs to be tested. Even assuming that response is centrally-mediated, there may be differences in response timing due to variations between body sites in skin permeation and absorption. These are likely to vary according to skin thickness and pigmentation, due to changes in structure of the stratum corneum and in cellular make-up and matrix components in the epidermis. There is limited data regarding influence of skin type on response to cold, and, although presence of hair follicles may change response timing, it is uncertain as to whether total
intensity or quality varies between glabrous and hairy skin (Davis, 1998; Harrison & Davis, 1999). This clearly needs to be evaluated with the menthol test.

Different skin pigmentation and exposure to sun also influences stratum corneum structure and so comparisons in response to menthol between individuals with different skin colour would clarify its significance. Finally, sensory response to cold may be influenced by climatic heritage. Although only a poor association has been found between response to environmental cold and response to a focal area of evoked cold, a difference in ADI score might logically be anticipated depending on whether an individual came from a cold northern hemisphere country or from a tropical country. For the most widespread applicability of the test, this factor may need investigation.

**Clinical implications**

The medical literature is increasingly concerned with ways to manage the growing number of individuals presenting with intractable chronic pain and response to cold is receiving greater acknowledgement as a potential indicator of altered pain processing associated with pain chronicity and severity. However, current approaches to the assessment of cold hyperalgesia, in a research context, let alone a clinical context, have severe limitations in terms of equipment cost, reliability or concerns about validity. Consequently cold hyperalgesia is often omitted from chronic pain studies and dismissed as either too variable or too difficult to assess.

The current study however has taken a non-thermal approach to the assessment of cold sensation and found that the topical menthol test may provide a clinically-applicable alternative to standard cold-temperature assessment. Within the limitations acknowledged above, this series of studies has demonstrated that a topical menthol formulation is a valid and reliable method for evaluating widespread cold hyperalgesia. In patients with chronic pain from knee OA, a menthol-cold hyperalgesic response at a site distant to the OA knee was shown to be associated with higher self-reported pain and dysfunction, greater actual dysfunction, greater extent of widespread mechanical hyperalgesia and higher report of significant levels of neuropathic-type pain. Menthol-cold hyperalgesia was therefore indicative of a combination of factors, which point to disabling persistent pain that is centrally-driven and augmented.

Whilst a simple test such as the ice application test (Maxwell & Sterling, 2012) is accessible and inexpensive for a clinical setting, it suffers from lack of precision in
measurement due to its unavoidable reliance on a supra-threshold stimulus. Menthol offers the advantage of being able to avoid the potential physiological and psychological complications of a thermal-cold stimulus by activating only a sensory response. Menthol concentration can also be varied so that a stimulus which is sub-threshold for the majority of individuals can be applied and so both hyperalgesic and a truely allodynic (abnormally noxious response to a non-noxious stimulus) responses can be measured. The more multi-dimensional ADI response measurement score also enables greater precision and sensitivity in measurement. Whereas ice application is able to identify extreme responses, the menthol ADI test has the potential to be able to differentiate between levels of pain augmentation. Small changes in intensity of cold or in type of sensation experienced are registered as changes in total ADI score, or more specific analysis of descriptors may show changes in selection of words, for example from burning to hot or from pickling to tingling, that indicate changes in nociceptive drive.

Although still speculative, this level of sensitivity in the ADI scoring system may mean that a clinician would be able to identify the degree of central augmentation experienced by an individual and so be able to select the most appropriate intervention based on pain (mechanism) rather than patho-etiolog. The menthol test may also be used to monitor efficacy of an intervention so that an ineffective medication can be changed at an earlier stage. However, further studies are needed to confirm this potential: for example, applying the menthol ADI test alongside additional QST measures of central excitability, spinal inhibitory dysfunction, imaging and biochemical analysis to confirm the association between different ADI scores and predominance of peripheral, central sensitisation or central augmentation. Additional studies are also needed to review whether high ADI score predicts persistent pain after surgery since modulation of centrally-driven pain pre-operatively may provide better post-operative results.

From a practical perspective, the menthol test is inexpensive and requires no particular skill to apply, aside from adherence to a standardised protocol. The test also is relatively quick, although as mentioned above, further investigation of the reliability of reducing application time from 15 to 10 minutes is still needed in order to improve practicality. There is no complex technology to be managed and its interpretation can be simple (ADI≥5) or more complex, involving analysis of individual intensity and
quality components. It is also likely to be more acceptable for the majority of individuals than an ice test.

Although there is considerable work still to do, if the initial findings of this current investigation can be confirmed and mechanistic hypotheses tested, a specifically designed topical menthol test with an appropriate scoring system such as the ADI has enormous clinical and research potential.
Chapter 11

Summary

This series of studies investigated the validity, reliability and clinical value of a topical menthol test quantified with a specifically designed measurement tool, for identifying cold hyperalgesia in individuals at risk of developing centrally-augmented pain.

This is the first study to demonstrate systematically that topical menthol evokes a concentration-dependent response. Clearly differentiated responses were seen in healthy subjects, with lower concentrations evoking predominantly cool or cold sensations and higher concentrations evoking unpleasantness and selection of noxious descriptors such as icy, burning or stinging. Remarkably similar response characteristics were also seen when the ADI scoring system was applied to sustained thermal-cold temperatures, with 20°C evoking innocuous cool sensations and 10°C evoking unpleasantness, pain and burning sensations. These sensory responses also provide valuable psychophysical support for basic science findings. The consistency with which menthol at a particular concentration and cold at a particular temperature caused a transition from innocuous to noxious sensations suggests similar signalling mechanisms for cold pain that may be related to change in fibre-type activation and altered inhibitory balance. By demonstrating the ongoing experience of cold alongside noxious sensations, this study also provided additional support for the debated role of TRPM8 in cold pain transduction. Importantly for the validity of the menthol tests these studies showed that, despite differences in precise transduction mechanisms at the TRPM8 channel, menthol at the appropriate concentration mimics the effects of cold temperature on the neurophysiological system without changing skin temperature and so may be valuable as a surrogate test stimulus.

Assessment of cold hyperalgesia using conventional cold pain threshold testing has been widely criticised as unreliable due to the wide variability in temperature values and difficulties with equipment. In Chapter 6, the menthol ADI test demonstrated similarly high reliability to CPT across two test sessions but this was associated with overall lower variability. This appears to relate to the use of a weighted, bi-dimensional measurement system for the ADI which is able to account for both the intensity ('volume') of pain system response and subtle pathway changes in the system when the
‘gear’ shifts from innocuous to mildly noxious quality of sensation. This suggests that the menthol test may be preferable for identifying cold hyperalgesia.

Since normal response to a particular concentration of menthol could comprehensively be characterised with intensity and quality ratings, the menthol test demonstrated a clear ability to differentiate between normal and abnormal responses in both healthy and clinical subject groups. The degree of normality/abnormality of response associated consistently well with conventional thermode CPT assessments, thereby showing good concurrent criterion validity. Within the same cold modality (Chapter 5) the ADI showed particularly good sensitivity and specificity in predicting whether an individual would be allocated to a high or low CPT group. In Chapter 7, using the cut-off score of ≥5 indicated by Study 3, the ADI identified individuals with knee OA who exhibited widespread cold hyperalgesia with similar accuracy to CPT.

Chapter 8 investigated whether widespread cold hyperalgesia, as defined by forearm menthol ADI ≥5, was able to identify a sub-group of individuals with knee OA with associated signs of centrally augmented pain. This study showed that grouping according to ADI≥5 differentiated the 24% of individuals who also exhibited widespread mechanical and heat hyperalgesia, CPT >15 at every test site, significant self-reported signs of neuropathic-type pain and higher levels of pain and dysfunction. This combination of factors associated with ADI-cold hyperalgesia is strongly suggestive of centrally-driven pain augmentation that may not be responsive to standard locally-acting anti-inflammatories or analgesics. Chapter 9 tested this hypothesis by investigating whether the Cox-2 inhibitor etoricoxib influenced only local measures of inflammatory hyperalgesia such as knee PPT and HPT or whether measures more reflective of centrally-driven pain mechanisms were also changed. Although the greatest overall effect of etoricoxib was found locally, etoricoxib did reduce both ADI and PainDETECT scores in many OA subjects. Although CPT temperature remained fixed, sensation quality at CPT did show an 11% reduction in burning sensation, which mirrored the main quality change exhibited by menthol ADI at the forearm. This indicates that the menthol ADI test, applied at a site distant from the painful area, is more sensitive to small changes in the pain system than conventional CPT temperature. If future studies are able to confirm these findings, the menthol ADI test has potential to enable stratification of patients according to degree of peripheral sensitisation, central sensitisation or central pain augmentation and facilitate provision of the most appropriate intervention before secondary effects of
chronic pain have become established. For example, a future line of testing would be to identify individuals showing signs of central pain augmentation, using the ADI and additional QST tests, and apply a blinded placebo or comparator-controlled design to evaluate the efficacy of a neuropathic pain medication such as pregabalin (Lyrica). The sensitivity of the test mean that it also has the potential to monitor the efficacy of a selected intervention by identifying subtle changes in the quality or intensity of widespread pain response.

This investigation has shown that the menthol ADI test is a reliable and valid addition to the QST tool-box with great potential clinical value, that may also encourage more extensive research into cold hyperalgesia as a chronic pain phenomenon.
Chapter 12

References


331


335


Chapter 12

References


Chapter 12

References


348


Chapter 12

References


350
Chapter 12


355


---


---

References


357


Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.
Appendices
Appendices

Table of Contents

Appendix 1: Development of a prototype menthol test - Liquid formulation
  1.1 Introduction ........................................................................................................365
  1.2 Development of a simple liquid formulation .....................................................366
  1.3 - Determination of the rate and extent of menthol release from the formulation
        .........................................................................................................................368
  1.4 - Development of a simple topical delivery method ..........................................369
  1.5 - Initial development of an initial sensory measurement tool ..........................371
  1.6 - Summary of Appendix 1 ..................................................................................372

Appendix 2: Refinement of the sensory measurement tool
  2.1 Development of a electronic VAS system .......................................................375
  2.2 Development of a refined descriptor index (ADI) .............................................376

Appendix 3: Additional test-retest data from Study 7 .................................................383

Appendix 4: Supplementary Data for Study 7 (Etoricoxib in Knee OA) ............387

Appendix 5: Further details of publications referenced in this Thesis .................391
Appendix 1

Development of a prototype menthol test

1.1 Introduction
The first step in the test development process involved the design and production of a simple preliminary prototype. This initial development of the menthol test involved laboratory studies and small pilot psychophysical investigations of four inter-linked aspects:

- Development of a simple formulation;
- Determination of the rate and extent of menthol release from the formulation;
- Development of a simple topical delivery method;
- Development of an initial sensory measurement tool.

Once a simple liquid solvent vehicle had been formulated, in vitro release of three concentrations of menthol across a semi-permeable membrane similar to skin was determined. Basic delivery issues were assessed in order to determine a suitable delivery/application protocol for Study 1 on human volunteers. Delivery aspects assessed included (i) quantity of formulation to be applied, (ii) anatomical site and surface area of skin, and (iii) potential value of occluding the formulation on the skin site.

The following section will only include basic aspects of this development due to Intellectual Property constraints. Additional details may be found by consulting the patents associated with this development (International Patent Application PCT/AU2009/001037: Method and Device for Determining the Severity of a Pain Disorder).

All laboratory preparation and testing was performed in the Skin Research Laboratory in the School of Pharmacy.
1.2 Development of a simple liquid formulation

The formulation development process was informed by an initial review of the literature in which topical menthol had been used as an experimental pain model and evaluation of the physicochemical properties of menthol and potential vehicles. Solubility studies were undertaken to determine the saturation solubility of menthol in a range of chosen vehicles. Based on these data a number of concentrations of menthol were prepared in the chosen vehicle and used for small psychophysical pilot studies (n=5). These determined the concentrations of menthol that might be suitable to discriminate between normal and abnormal responses to cold. Based on these studies, three concentrations of menthol in the vehicle were then used in the laboratory release studies.

Physico-chemical properties of menthol

Menthol is a natural cyclic terpene derived from plants in the Mentha species and has been incorporated extensively in pharmaceuticals, confectionary, digestive and oral hygiene products. Menthol has a hexane ring with three asymmetric carbon atoms which generate four pairs of optical isomers, however, only levo-menthol (l-menthol), as shown in Figure 1 exerts a cooling sensation when applied to the skin. At room temperature and in its basic form, l-menthol is a colourless solid substance with a needle-like crystalline structure.

![Figure A1.1. Chemical structure of Levo-Menthol](image)

Menthol has suitable physiochemical characteristics for optimal permeation of the outermost stratum corneum region of the epidermis, which is the main barrier to skin penetration of topically applied substances. These include its small molecular size and lipophilic character. In addition terpenes such as menthol have a direct penetration enhancing effect on the stratum corneum by disrupting the hexagonal hydrocarbon chains of the lipid domains thus reducing their barrier properties (Obata et al., 2006).
Appendix 1

As a result menthol is often used as a penetration enhancer to enable more efficient topical absorption of other pharmaceutical agents. Thus it is expected that menthol will rapidly and effectively permeate the skin.

**Solvents**

For topical application, menthol needs to be in solution in order to make it available for skin permeation, thus a suitable solvent is required. Menthol is lipophilic and thus almost insoluble in water but highly soluble in organic solvents such as ethanol, ether and chloroform. Propylene glycol enhances the solubility of menthol in an aqueous solution and is a commonly used cosolvent in formulation vehicles. A review of the literature provides many reports in which menthol has been included in a topical formulation as a penetration enhancer for another active substance, but the permeation of the menthol itself is not determined. In these reports ethanol is the most frequently utilized vehicle component to enhance solubility and has the advantage of also contributing to enhanced skin permeation. For example, Fang et al. (2008) found that a combination of ethanol and menthol enhanced permeation of tetracaine gel more than twice as effectively as each substance on its own. The optimal concentration for ethanol as a transdermal penetration enhancer was 70%w/w in water in that study (Fang et al., 2008).

Tween 80 has also been considered as an additional vehicle constituent. Tween-80 is a non-ionic surfactant often used in pharmaceutical formulations as a wetting agent for poorly aqueous soluble drugs. It reduces the interfacial tension between the particle surface and the solvent thus facilitating the dissolution process. Hatem et al. (2006) reported adding 1% Tween-80 to the aqueous-ethanol vehicle in their topical menthol preparation but did not assess its benefit. Kopec et al. (2008) suggested that Tween 80 was of limited effect in improving solubility of capsaicin in an ethanol-saline vehicle.

Small-scale solubility studies were carried out as a preliminary to the main studies to investigate whether Tween 80 was a necessary addition for solubility purposes.
1.3. **Determination of the rate and extent of menthol release from the formulation**

The release of menthol from vehicle solutions was determined using an ILC14 Automated System with 14 In-Line Cells (Figure A1.2). The release of different concentrations of menthol through a semi-permeable membrane was determined. The amount of menthol in all samples was determined by GC Flame Ionization Detection (FID).

**Key questions included:**

- What are the release characteristics of menthol from solution across a semi-permeable membrane?
- Is there a relationship between amount of menthol release and concentration?
- Is there a relationship between rate of menthol release (time to onset or time to peak) and concentration?

The Permagear In-Line Cell Automated system has 14 cells consisting of a donor compartment and a flow-through receptor compartment that is connected via plastic tubing to collection cells in a fraction collector. The donor and receptor compartments were separated by a pre-wetted cellulose semi-permeable membrane. The receptor solution of 50% aqueous ethanol constantly flowed through the receptor cells at a rate of 55μl/min to ensure sink conditions for diffusion of menthol. Receptor solution fractions were collected over 10-min for a total experimental period of 1h. The in-line cells and receptor solution were maintained at 35°C to simulate body temperature and all cells were sealed with Parafilm® to limit evaporation of ethanol.

![Figure A1.2](image_url)

*Figure A1.2*  
Permagear ILC14 Automated System with 14 In-Line Cells
The same process was followed for each of the menthol solutions. Menthol demonstrated statistically significant dose-dependent release.

1.4 Development of a simple topical delivery method

The goal was to develop a reliable and consistent procedure to allow meaningful assessment of the initial formulation during the first Study (Chapter 3). The optimum approach to measurement of sensory response will be considered in section 4.2.4 below.

Key Questions:

• What is the best method to enable consistent contact between the liquid formulation and skin during application?
• Does the formulation need to be occluded?
• What quantity of formulation is needed, across what area of skin?
• Does skin type make a difference (glabrous versus hairy)?
• What is the likely time-frame to achieve peak response during initial testing?

The protocols used in previous studies involving topical menthol in psychophysical testing were reviewed. The subsequent concepts were then evaluated with small pilot studies in four or five subjects before a prototype test procedure was finalised.

Seven previous studies have applied menthol topically to human subjects. Table 1.3 lists the delivery and sensation measurement parameters used by 6 of these studies. The 7th study Yosipovitch et al. (1995) provided little information about exact delivery or measurement methodology so could not be included. The majority of these studies were in some way inter-related – for example, Hatem et al. (2006) explicitly copied the
Appendix 1

methodology used by Wasner et al. (2004) with 30% rather than 40%w/w menthol. Consequently it is not possible to discern whether the reported parameters were selected for best efficacy or for best comparability with previous studies. All studies applied a liquid aliquot of menthol in an aqueous-ethanol solvent. This solution was applied to a pad of some kind (gauze, cellulose or filter paper) under occlusion to prevent evaporation of the ethanol vehicle, although using a variety of materials including cling-film, a wide strip of parafilm, adhesive film and a silicon chamber (no other details provided). Several studies included an additional means to improve skin contact between the soaked pad and the skin such as a wide rubber band. The area to which the solvent was applied varied between 12.5 cm² to 16 cm², yet the quantity of solvent was between 1 and 2 ml, with no rationale provided for area or quantity used in most cases. The majority of studies used glabrous skin on the volar forearm. Studies using menthol as a sensitizer in clinical populations did not state a test site so it is assumed that this was governed by the site of pathology. Menthol application time varied between 10 and 20 min. Again no rationale was provided for this choice.

In our initial evaluations we determined that using cling-film to occlude the solution was unsuitable as the ethanol started to evaporate and create a cold sensation. Parafilm and adhesive film were also found to be unsuitable. Tegaderm® clear adhesive dressing (3M), was selected since it was easy to apply quickly once the liquid had been added to the gauze, was easy to apply across a very defined area due to the paper ‘window’ provided, and the adhesive is firm enough to limit leakage even with arm movement.

![Figure A1.4: Tegaderm® clear adhesive dressings](image)

Although ultimately a 10-min time-frame for test application in a clinical setting would be ideal, at this developmental stage the sensory response over a longer time period was assessed to provide more rich and meaningful data. It was therefore decided to apply the menthol for 15-min but continue to assess response for an additional 10-min.
1.5 Development of an initial sensory measurement tool

The aim of the menthol test is to identify cold hyperalgesia. Any test intended to elicit a sensory response, which can then be classified as normal, or hyperalgesic, requires a measurement tool with which to quantify the response.

The menthol test is intended to be an alternative method for assessing cold hyperalgesia, a phenomenon conventionally quantified by cold pain threshold. There are fundamental differences between CPT testing and the proposed menthol test. Aside from differences in the nature of the stimulus there are differences in the way in which the stimuli are delivered and the type of response obtained, which will need to be reconciled when developing the menthol measurement tool. During CPT testing an individual is provided with a particular sensation on which to focus (“painful cold”), is then provided with a descending temperature stimulus and asked to select the temperature that matches. In contrast to CPT, the menthol test follows a QST approach whereby the individual is provided with a standardized stimulus and then asked to provide their response to that stimulus. Whilst CPT only provides information about elevated intensity within an anticipated sensory framework, it would be advantageous if the menthol test is able to measure both changes in intensity and quality of sensation.

Intensity Measurement

The simplest and best-established approach to measurement of intensity is through the application of visual analogue scales (VAS). Although VAS scales are most often used to measure intensity of pain, they have also been applied to other sensory phenomena such as breathlessness.

It was therefore proposed that normal and abnormal sensory responses to the menthol stimulus should be measured initially with VAS scales for the key sensations of cold, unpleasantness and pain. It might be anticipated that if menthol is activating TRPM8 the primary normal sensation would be that of cold, increasing in intensity with higher concentrations. An abnormal response on the other hand, is likely to involve some degree of pain or discomfort at a normally non-noxious concentration. It was decided therefore to include VAS scales for pain and unpleasantness as well as cold.
Appendix 1

**Quality measurement**

Several previous studies measured quality of sensation in addition to intensity. The advantage of quality data is that it offers a broader characterization of the sensory experience. When characterizing normal and abnormal responses, a more comprehensive understanding of not just the intensity but also the type of sensation experienced may provide better insight into the mechanisms driving the abnormal response. For example, previous menthol studies using descriptors have reported selection of the nociceptive word ‘burning’ by high proportions of subjects.

Although Hatem et al (2006) allowed subjects completely free choice of word(s) to describe their sensation in response to menthol, it is proposed that in the current study a more constrained, although still comprehensive, list of words will be provided from which subjects can choose those most appropriate to describe their sensory experience. Previous studies (Wasner 2004, Binder 2011) have used the McGill Pain Questionnaire (MPQ) descriptor list. Although lengthy (70 words), it is a well-established and validated list that includes sensory as well as affective and emotive words. The MPQ was therefore used to characterize the sensory qualities induced by menthol stimulation.

**Summary**

The initial sensory measurement will include three VAS scales for cold, unpleasantness and pain (end-point descriptors used). Subjects will be asked to rate the intensity of each of these sensations at 1-min intervals throughout the application. Area under the VAS Vs time curve and maximum VAS intensity will be calculated for analysis. Subjects will also be asked to select as few or as many words from the MPQ pain descriptor list to describe the sensation they are experiencing (Melzack, 1975). Descriptor choice will initially be analysed by calculating the McGill Pain Rating Index (a score comprising the sum of weights for each word chosen) and the Number of Words Chosen (NWC index).
### Appendix 1

**Table A1.1:** Delivery and sensory parameters recorded by previous studies using topical menthol

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delivery Parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Formulation</strong></td>
<td>40% menthol</td>
<td>40% menthol</td>
<td>40% menthol</td>
<td>30% menthol</td>
<td>10% menthol</td>
<td>40% menthol</td>
</tr>
<tr>
<td></td>
<td>90% ethanol</td>
<td>90% ethanol</td>
<td>90% ethanol</td>
<td>90% ethanol</td>
<td>95% ethanol</td>
<td>90% ethanol</td>
</tr>
<tr>
<td><strong>Application area</strong></td>
<td>Gauze pad</td>
<td>Gauze pad</td>
<td>Gauze pad</td>
<td>Gauze pad</td>
<td>Filter paper</td>
<td>Cellulose pad</td>
</tr>
<tr>
<td></td>
<td>2.5 x 5cm</td>
<td>3 x 3cm</td>
<td>3 x 3cm</td>
<td>2.5 x 5cm</td>
<td>4 x 4cm</td>
<td>2.5 x 5cm</td>
</tr>
<tr>
<td><strong>Occlusion?</strong></td>
<td>Adhesive film + rubber band 2.5cm wide</td>
<td>Adhesive film + rubber band 2.5cm wide</td>
<td>Adhesive film + rubber band 2.5cm wide</td>
<td>Cling-film</td>
<td>Parafilm strip weighted at either end</td>
<td>Silicon chamber</td>
</tr>
<tr>
<td><strong>Quantity</strong></td>
<td>1ml</td>
<td>1ml</td>
<td>1ml</td>
<td>2ml</td>
<td>Not reported</td>
<td>200 µL</td>
</tr>
<tr>
<td><strong>Test site (skin type)</strong></td>
<td>Volar forearm (glabrous)</td>
<td>Variable (patient cohort) (hairy)</td>
<td>Hand dorsum (glabrous)</td>
<td>Volar forearm (glabrous)</td>
<td>Volar forearm (glabrous)</td>
<td>Variable (patient cohort)</td>
</tr>
<tr>
<td><strong>Application Time</strong></td>
<td>20 mins</td>
<td>15 mins</td>
<td>20 mins</td>
<td>10 mins</td>
<td>15 mins</td>
<td>20 mins</td>
</tr>
<tr>
<td><strong>Sensory Parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>How often measured?</strong></td>
<td>Once at end of application</td>
<td>N/A</td>
<td>NRS every 1 min McGill every 5 mins</td>
<td>Once at end of application</td>
<td>Once at end of application</td>
<td>Once at end of application</td>
</tr>
<tr>
<td><strong>Intensity</strong></td>
<td>NRS Pain 0-10</td>
<td>NRS pain 0-10</td>
<td></td>
<td></td>
<td></td>
<td>NRS Pain 0-10</td>
</tr>
<tr>
<td><strong>Temperature rating:</strong></td>
<td>Not assessed</td>
<td>Temperature rating: -100 (max cool to +100 (max warm) Not assessed</td>
<td>Temperature rating: -100 (max cool to +100 (max warm) Not assessed</td>
<td>Labelled Magnitude Scale (category ratio scale) for thermal &amp; nociceptive</td>
<td>Cold 0-10</td>
<td>Itch 0-10</td>
</tr>
</tbody>
</table>
**Appendix 1**

<table>
<thead>
<tr>
<th>Quality</th>
<th>McGill descriptor for pain quality</th>
<th>McGill descriptor for pain quality</th>
<th>Open: describe quality of sensation</th>
<th>sensations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not assessed</td>
<td>Not assessed</td>
<td></td>
<td>Not assessed</td>
</tr>
</tbody>
</table>

* These three studies were carried out by the same research group explicitly using the same methodology
Appendix 2

Refinement of the sensory measurement tool

2.1 Development of a electronic VAS system

Studies 1 and 2 clearly showed that the paper version of the four VAS scales was problematic from an administrative perspective. Paper VAS scales were awkward for the subject to complete, requiring a new sheet every one minute. They were also time-consuming to measure and convert manually to scores. Both of these factors would make clinical application of the tool unwieldy and inefficient. It was therefore decided to develop an electronic VAS system.

A number of previous studies have used electronic VAS systems and found them to be equally reliable and valid as paper versions (Wright et al., 1990; Jamison et al., 2002). Although a touch-screen application was the preferred option, financial constraints in the current investigation meant that a less sophisticated system needed to be developed for interim use. Four linear potentiometers with an attached slider, were encased in a metal box. Each slider was labelled with end-points “maximum ...” and “minimum ...” for each of the VAS sensations (cold, heat, unpleasantness and pain). The box was designed to limit the physical travel of the potentiometers to 10 cm. The potentiometers were linked to LabVIEW (NIC, Austen, USA) software which calibrated the sliding scales to 100mm and recorded the values set by the subject at each minute. A timing signal within the LabView program activated a red light every one minute to prompt recording of VAS values.

![Figure A2.1: e-Vas Box](image-url)
Appendix 2

The new e-VAS was piloted in a small group of five pain-free subjects. Gel B was applied on two occasions, separated by 24 hours, to the volar forearm of each subject for 10 minutes using the same methodology as in Preliminary Study 2. On each occasion the subject rated their intensity of cold, unpleasantness, heat and pain each minute, either with the paper VAS sheets or e-VAS. Order of VAS medium was randomly allocated. Subjects also provided a single descriptor rating at the end of each allocation. This was to assess comparability of sensation experienced on each occasion. Results showed that there was no significant difference between PRI score for descriptors between sessions. There was also no significant difference in AUC VAS value between paper and e-VAS media for any of the sensations (cold p=.245; heat p=.487; unpleasantness p=.158; pain p=.779). Correlation coefficients were not run since study numbers were so small. Although only a very small pilot study, this suggests that there is reasonable comparability between paper administered VAS scales and electronic scales.

2.2 Development of a refined descriptor index (ADI)

Background
The goal of the menthol test is to be able to identify individuals with abnormally elevated responses to cold stimuli (cold hyperalgesia). Whereas the CPT approach quantifies cold hyperalgesia with a temperature, menthol requires development of an index which allows sensory features to be ranked in such a way that a high score reflects a cold hyperalgesic response. This response must also show some association with CPT values for cold hyperalgesia.

There are existing scales which seek to assess pain through quantification of quality descriptors. Self-report scales such as the McGill Pain Questionnaire (MPQ) (Melzack, 1975) and the more recent Pain Descriptor Scale (Fernandez et al., 2011; Fernandez et al., 2012) assign a weighted value to a series of categorised words. In both cases, words are rank-ordered based on perceived severity within a sub-category, as agreed by doctors and chronic pain patients. These scales do not aim to diagnose the cause of pain but aim instead to provide more comprehensive information to guide pain management.

A large number of other instruments use pain descriptors to diagnose the cause of pain. In particular there has been a recent surge in self-report scales which assess the presence of pain qualities commonly associated with neuropathic pain: e.g. Pain Quality Assessment Scale (PQAS), PainDETECT and Doleur Neuropathique 4 (DN4).
Appendix 2

(Freynhagen et al., 2006; Victor et al., 2008; Bouhassira & Attal, 2011). These instruments do not assign rankings to words but ask an individual about the presence of specific neuropathic-pain type qualities. Scores above a cut-off indicate presence of neuropathic pain. The Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) weights the answers to five questions about presence of neuropathic symptoms with two physical tests of sensory loss or augmentation to provide a similar cut-off score to diagnose neuropathy (Bennett, 2001).

Therefore the existing pain quality tools either have a very specific diagnostic focus or a broader mandate than is intended for the menthol test. These pre-existing instruments have also all been designed to quantify everyday pain qualities rather than the quality of evoked pain during QST, as is required for the current menthol test.

The following section explains in fairly general terms how the Algotect Descriptor Index was developed. Additional details about weightings and scoring development remain Commercial In Confidence for IP purposes (ref: International Patent Application PCT/AU2009/001037: Method and Device for Determining the Severity of a Pain Disorder).

Analysis of descriptor data in Studies 1 & 2

A starting point for assessment of the sensory effects of menthol was needed and so it was decided to use the MPQ Descriptors list. This list with its associated Pain Rating Index (PRI) is the most widely used and validated descriptor index, including for use in experimental pain (Klepac et al., 1981) so, although not a perfect fit for the current purpose, would provide a broad indication of whether the menthol test evoked more intense sensation quality in those with conventionally-measured cold hyperalgesia. It was predicted that the McGill PRI would only provide this initial evidence but that the descriptor data collected could then be analysed independently of the McGill rankings in order to evaluate whether selection of certain words was associated with other abnormal or cold hyperalgesic characteristics.

The PRI showed significant dose-dependent changes for both the liquid and gel formulations of menthol, indicating some association between intensity of sensation and choice of higher ranked pain descriptors. When subjects were divided according to CPT<>15°C, PRI was higher at all concentrations for cold hyperalgesic subjects but only reached significance at one concentration. Although the PRI score illustrated a
Appendix 2

relationship between CPT and higher ranked pain words, the association with the total index value was not sufficiently strong to clearly identify cold hyperalgesic individuals. This is not surprising since the choice of words and their rankings in the PRI are intended to characterise the day-to-day chronic pain experience, not identify individuals with a specific pain quality.

An alternative index was therefore designed using weightings for each word based on the frequency of its selection by subjects showing a cold hyperalgesic or clearly abnormal response. Individual word choice for Study 1 was first analysed to explore whether those with cold hyperalgesia (CPT>15°C) selected a particular sub-set of words. For the gel study, there was no CPT with which to compare reponses to menthol so cluster analysis was used to separate out a smaller group of subjects exhibiting a more intense response. The 30% of subjects in the high PRI cluster group could once again be differentiated by their more frequent selection of specific words, which correlated strongly with those chosen by the high CPT group in Study 1.

Even though these studies were small, the relative consistency of word selection by the minority high intensity or hyperalgesic groups corresponds with what might be proposed as the likely characteristics of cold hyperalgesia. A new word weighting system was therefore developed based on this word choice by the hyperalgesic groups. The index was named the “Mean Word Score” (MWS).

Intensity data

Measures of varying intensity are also likely to be components of an abnormally elevated response to cold. The final index measure therefore ideally needed to combine both quality and intensity data. In this way, the new menthol test would potentially offer improved content validity as a measure of cold hyperalgesia in comparison to the current CPT temperature.

However this immediately raised questions about how to incorporate VAS values. Should all VAS values be included? Should the values be incorporated as weighted or raw scores? Increasingly studies are suggesting that pain intensity and quality are of equal importance because they reflect different but complimentary pain domains. So, although determining suitable weighting scores for each VAS sensation would involve
an additional step away from raw values, the total index score would then better reflect
a relatively equal balance between intensity and quality values.

The relationship between hyperalgesic response and VAS ratings was less clear than
for descriptors. Both the liquid and gel study showed that an abnormal/hyperalgesic
response was associated with some report of pain and unpleasantness. It was unclear
as to whether a cut-off VAS value would be useful, particularly for unpleasantness since
for example during Gel B application almost 75% of subjects reported some degree of
unpleasantness, although 3% of these individuals rated that unpleasantness as very mild
- 20/100 or less. Heat VAS was also ambiguous. It had been anticipated that sensation
of paradoxical heat would be an important distinguishing feature for a cold
hyperalgesic response, thus its addition to the VAS scales for the gel study. However, it
was found that almost 50% of subjects reported some VAS heat sensation, agreeing
with previous studies which found a high incidence of warm sensations during
particular phases of thermal cold stimuli (Davis and Pope, 2002). Four subjects in the
menthol gel study reported higher heat and also recorded high levels of pain and/or
unpleasantness and high PRI scores. So it appears that heat VAS above, but not below, a
certain a cut-off value may distinguish abnormal response.

Based on this data, and applying weightings retrospectively to both the liquid and gel
studies, decisions were made about an initial weighting system for VAS scores.

**Algorect Descriptor Score**

A final index, known as the Algorect Descriptor Index (ADI) was therefore determined.
The index combined weightings for each VAS value (to a possible maximum of 4) plus
the score for descriptors (MWS sub-score: maximum possible 5). The total possible for
this new score was therefore 9. Although validity of the newly developed ADI and MWS
sub-scores was to be tested in Study 3, in the interim ADI calculations were carried out
retrospectively on the data from Study 2.

Fig A2.3a and b below show the distribution curves and box plots for ADI and PRI. ADI
scores showed better homogeneity and slightly more normal distribution, although still
with a trend towards higher scores.
Figures A2.3.a-c: Normal distribution curves for a) ADI and b) PRI scores during application of menthol gel. c) Boxplots for ADI and PRI scores for Gel B

The new ADI showed moderate to good correlations with raw VAS values for heat (r = .595, p = .001), unpleasantness (r = .480, p = .011) and pain (r = .584, p = .001) but not for cold (r = .124, p = .537). Associations between ADI and PRI were also explored, with a good positive correlation shown (Figure A2.4). Although this might seem self-evident, the ADI weighting system included VAS intensity scores as well as a descriptor sub-score. It is interesting to note that the MWS (new descriptor sub-score) showed no correlation with PRI (r = .249, p = .211) which reflects the differences between MWS and PRI in both word weightings and method of calculation (PRI uses the sum of weightings; MWS uses the mean weighting). Cluster analysis with ADI produced a cut-off value of 4.5-4.8/9. Significant differences were shown between high and low ADI scores for PRI (P = .008), heat VAS (p = .016), unpleasantness VAS (p = .018) pain VAS (p = .013) but not for cold VAS (p = .863), as might have been predicted. However, group size was not appropriate, with the majority of subjects (63%) clustered into the high ADI group. Subsequent analysis suggested that a higher cut-off value of 5 was more discriminative.
ADI and MWS sub-scores were fully evaluated using cold-temperature data (sustained and threshold stimuli) from Study 3. Study 3 investigated sensory response to sustained cold stimuli and compared this with concurrently tested CPT. PRI and ADI / MWS scores were compared and correlations between both indices and CPT explored in order to assess criterion validity. Results from Study 3 showed that the ADI was able to discriminate significantly between sustained thermal-cold temperatures. Both VAS and descriptor sub-components of the ADI score showed significant temperature-dependent differences and word choice was clearly more intense and unpleasant at the lowest temperature. When ADI scores were compared with CPT values, ROC curve analysis showed that ADI was a more accurate predictor of CPT group (<15°C) than PRI with a sensitivity of .90, and specificity .74. There was also a strong (non-statistical) relationship between quality of sensation experienced at 10°C and for Gel B, showing that cold temperature and menthol stimulus at similar intensities evoked similar responses with the ADI. This study therefore demonstrated that the ADI has good criterion validity and potentially better content validity than the standard CPT test measure due to its inclusion of several pain dimensions in the one score.

2.2 Summary
Several refinements were made to the sensory measurement tool. An electronic VAS system was developed to improve efficiency of VAS intensity data collection. A newly developed index was created that combined weighted VAS intensity data and sensory quality data into a single score: the Algotect Descriptor Index (ADI). Both of these developments were then tested during Study 3. Based on the results from Study 3 it was determined that the ADI provided a discriminative measure of cold hyperalgesia associated with menthol application. This provided the basis for the menthol test applied in the subsequent clinical studies.
Appendix 3

Additional test-retest data from Study 5

Background
Study 4 (Chapter 6) evaluated the test-retest reliability of the menthol ADI test with healthy subjects. This study found that, although ICC values were high, there was a pattern of variability such that, at Day 2, all VAS and MWS scores reduced, by 5-8% from baseline. Intensity of heat and choice of the words warm and hot were particularly reduced by Day 2. It was decided to run a similar ICC analysis on data from the Placebo group of Study 7 to assess whether the same pattern was shown. This group was washed out of usual analgesia for both test sessions and the time-frame between their initial session and first follow-up test at Day 4 was similar to the test-retest time-frame for Study 2. If a similar pattern of reduced values on the second test occasions is found in this study, there may be implications for scaling of ADI scores.

Method
The menthol ADI data from Day 0 and Day 4 for 40 subjects with knee OA was included in this analysis. These subjects were voluntarily recruited for Study 7 (Effect of etoricoxib in knee OA) according to the criteria listed in Chapter 9 and had been randomly allocated to the Placebo arm of the study. Subjects had been washed out of their usual NSAID and analgesic medication before Day 0 testing. Although all subjects were provided with acetaminophen prn as rescue medication, they were requested to refrain from taking any medication 12 hours before their test session on Day 4.

Menthol testing on Day 0 and Day 4 followed the procedure as described in Chapter 9, with data for VAS intensity for cold, heat, unpleasantness and pain and descriptor data collected at regular time-points during the 15-minute application of occluded menthol gel to the volar forearm. An ADI total score was calculated, as described in Appendix 2, composed of the combined weighted VAS values and descriptor MWS score.

Each of these values was analysed for reliability and variability using the same method as for Study 4, using Intra-Class Correlation Coefficients (ICC), including 95% confidence intervals (CI) and calculating Standard Error of Measurement (SEM) and Correlation Coefficients, illustrated with scatterplots. SPSS v19 statistical package was used and alpha set at p<.05.
Appendix 3

Results
The mean age of these 40 subjects was 65 years (range 50-86 years) and the group comprised 19 males and 21 females.

• ADI
Table A3.1 illustrates the ICC values for the total ADI and MWS descriptor and VAS intensity sub-components. The ADI showed high reliability (ICC=.917) and low 95% CI. There was other evidence of low variability: Figure A3.1 illustrates that there was very good correlation in ADI score between the two days (r=.847) and Figure A3.2 shows that the difference between the two test days was minimal (4% reduction by the second test day).

Table A3.1: Intra-Class Correlation Coefficient (95% confidence interval) for ADI total score and MWS and VAS sub-scores

<table>
<thead>
<tr>
<th></th>
<th>ICC</th>
<th>95% CI</th>
<th>Mean Difference (Day 1-Day2)</th>
<th>Standard Error of the Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADI total score</td>
<td>.917</td>
<td>.870 - .946</td>
<td>.159</td>
<td>.276</td>
</tr>
<tr>
<td>MWS score</td>
<td>.886</td>
<td>.823 - .927</td>
<td>.056</td>
<td>.172</td>
</tr>
<tr>
<td>VAS Cold</td>
<td>.889</td>
<td>.827 - .929</td>
<td>2.66</td>
<td>5.20</td>
</tr>
<tr>
<td>Heat</td>
<td>.674</td>
<td>.492 - .791</td>
<td>-2.96</td>
<td>8.22</td>
</tr>
<tr>
<td>Unpl</td>
<td>.873</td>
<td>.802 - .918</td>
<td>1.33</td>
<td>5.10</td>
</tr>
<tr>
<td>Pain</td>
<td>.897</td>
<td>.840 - .934</td>
<td>.065</td>
<td>2.47</td>
</tr>
</tbody>
</table>

Figure A3.1: Overlay scatterplot showing the strong correlation between Day 0 and Day 4 scores for total ADI score and MWS descriptor sub-score.
• **MWS descriptor sub-score**

MWS descriptor scores showed similarly good reliability (ICC = .886) and low 95% CI range (Table A3.1). Variability was slightly greater than for ADI but still very low: SEM .172 and Correlation Coefficient r = .796 (Figure A3.1).

![Figure A3.2](image)

**Figure A3.2:** Mean (SEM) values for ADI and MWS for Placebo OA subjects on Day 0 and Day 14

• **VAS intensity values**

VAS intensity values for cold, unpleasantness and pain showed similar reliability values to MWS sub-score and similar 95% CI ranges, with none below ICC= .802 (Table A3.1). Although cold and unpleasantness VAS were both lower on Day 2, this was by less than 10% in both cases (Figure A3.3). Day to day correlations were moderate to high for each VAS sensation: cold r = .800; unpleasantness r = .776; pain r = .814 (Figures A4.4a and b).

![Figure A3.3](image)

**Figure A3.3:** Mean (SEM) values for VAS sub-scores for Placebo OA subjects on Day 0 and Day 14

Heat VAS however showed considerably lower reliability and greater variability, with an ICC of .674 and wide 95% CI (.492 to .791). VAS values increased by 60% by Day 2
and mean difference showed a relatively large SEM (8.22). Compared with other VAS and MWS values heat showed a low correlation between days: r = .543 (Figure A3.4a).

**Figures A3.4a-b:** Overlay scatterplots showing correlations between Day 0 and Day 4 for VAS intensity values for a) cold and heat; b) unpleasantness and pain.

**Discussion: Key Points**

This study analysed the test-retest reliability of the ADI in a cohort of subjects with knee OA who were washed out of their usual analgesia on both test occasions.

Reliability statistics for ADI, MWS and VAS cold, unpleasantness and pain were overall good, but heat VAS showed considerably lower reliability.

Heat, unpleasantness and pain VAS increased by Day 4: VAS increased by 9% for pain, 15% for unpleasantness and 37% for heat. In contrast, cold VAS stayed much the same and MWS descriptor sub-score changed very little (4% increase).

When compared with results from the normative test-retest study in Chapter 6, although overall reliability was similar, the OA group showed a different pattern of change by Day 2. Whereas VAS values for the healthy group tended to slightly reduce on Day 2, values for the clinical group showed a greater tendency to increase by Day 2. In particular heat increased by 37% by Day 2 in the OA cohort.

This result suggests that further test-retest studies are needed to clarify whether there are differences in reliability between normal and healthy cohorts and whether this has implications for the ADI scoring system. In the Placebo analysis, an increase in heat and unpleasantness is most likely explained by increasing inflammatory symptoms due to lack of anti-inflammatory or analgesic medication.
Appendix 4

Supplementary Data for Study 7 (Etoricoxib in Knee OA)

Table A4.1: Placebo and Active group Day 0 mean scores for pain, dysfunction and quality of life.

<table>
<thead>
<tr>
<th></th>
<th>Placebo (n=40)</th>
<th>Active (n=40)</th>
<th>t(78)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WOMAC scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain (/50)</td>
<td>16.2 (9.5)</td>
<td>16.0 (7.7)</td>
<td>.093</td>
<td>.926</td>
</tr>
<tr>
<td>Stiffness (/20)</td>
<td>8.2 (4.7)</td>
<td>8.2 (4.3)</td>
<td>.017</td>
<td>.986</td>
</tr>
<tr>
<td>Function (/170)</td>
<td>55.4 (29.9)</td>
<td>51.5 (31.3)</td>
<td>.576</td>
<td>.566</td>
</tr>
<tr>
<td><strong>Pain intensity (VAS)</strong> (last 7 days)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (range) /10</td>
<td>4.5 (3-7)</td>
<td>4.6 (3-8)</td>
<td>-.190</td>
<td>.850</td>
</tr>
<tr>
<td>Strongest (range) /10</td>
<td>6.5 (3-10)</td>
<td>6.6 (3-9)</td>
<td>-.209</td>
<td>.835</td>
</tr>
<tr>
<td><strong>PainDETECT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score /35 (SD)</td>
<td>11.9 (5.3)</td>
<td>12.0 (5.3)</td>
<td>-.063</td>
<td>.950</td>
</tr>
<tr>
<td>&lt;13 (&quot;negative&quot;) (n)</td>
<td>20</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13-18 (&quot;unclear&quot;) (n)</td>
<td>15</td>
<td>12</td>
<td>χ²=.519</td>
<td>.771</td>
</tr>
<tr>
<td>19+ (&quot;positive&quot;) (n)</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PQAS sub-scores:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paroxysmal (/40)</td>
<td>9.9 (7.9)</td>
<td>9.9 (6.0)</td>
<td>.000</td>
<td>.999</td>
</tr>
<tr>
<td>Surface (/60)</td>
<td>8.2 (6.9)</td>
<td>10.3 (8.7)</td>
<td>-1.22</td>
<td>.226</td>
</tr>
<tr>
<td>Deep (/50)</td>
<td>15.2 (8.1)</td>
<td>17.0 (9.3)</td>
<td>.934</td>
<td>.353</td>
</tr>
<tr>
<td><strong>SF-36 Quality of Life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical sub-score</td>
<td>38.7 (7.5)</td>
<td>39.9 (7.7)</td>
<td>-.662</td>
<td>.510</td>
</tr>
<tr>
<td>Mental sub-score</td>
<td>54.3 (8.5)</td>
<td>51.8 (10.6)</td>
<td>1.15</td>
<td>.253</td>
</tr>
</tbody>
</table>
Table A4.2: Baseline physical function and quantitative sensory test values for Placebo and Active groups

<table>
<thead>
<tr>
<th></th>
<th>Placebo (n=40)</th>
<th>Active (n=40)</th>
<th>t(78)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALFS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (secs)</td>
<td>36.3 (13.1)</td>
<td>39.2 (15.7)</td>
<td>-.884</td>
<td>.379</td>
</tr>
<tr>
<td>3m Chair transfer (secs)</td>
<td>6.4 (2.0)</td>
<td>6.9 (2.3)</td>
<td>-.111</td>
<td>.273</td>
</tr>
<tr>
<td>8m return walk (secs)</td>
<td>12.6 (3.4)</td>
<td>13.2 (3.9)</td>
<td>-.816</td>
<td>.417</td>
</tr>
<tr>
<td>Stairs (secs)</td>
<td>17.3 (8.5)</td>
<td>19.0 (10.1)</td>
<td>-.818</td>
<td>.416</td>
</tr>
<tr>
<td><strong>QST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pressure Pain Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index knee</td>
<td>268.3 (135)</td>
<td>283.2 (117)</td>
<td>-.527</td>
<td>.600</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>325.8 (146)</td>
<td>330.6 (124)</td>
<td>-.158</td>
<td>.875</td>
</tr>
<tr>
<td>Elbow</td>
<td>331.5 (125)</td>
<td>350.1 (130)</td>
<td>-.612</td>
<td>.542</td>
</tr>
<tr>
<td>Cold Pain Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index knee</td>
<td>12.4 (8.9)</td>
<td>11.4 (9.6)</td>
<td>.497</td>
<td>.621</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>10.8 (9.6)</td>
<td>8.8 (8.3)</td>
<td>.044</td>
<td>.313</td>
</tr>
<tr>
<td>Elbow</td>
<td>10.3 (8.1)</td>
<td>9.8 (8.7)</td>
<td>.670</td>
<td>.820</td>
</tr>
<tr>
<td>Heat Pain Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index knee</td>
<td>44.8 (3.0)</td>
<td>45.8 (3.1)</td>
<td>-.143</td>
<td>.158</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>45.2 (3.2)</td>
<td>45.8 (2.8)</td>
<td>-.988</td>
<td>.326</td>
</tr>
<tr>
<td>Elbow</td>
<td>45.5 (4.1)</td>
<td>47.3 (3.4)</td>
<td>-2.11</td>
<td>.038*</td>
</tr>
<tr>
<td>Punctate Pain Threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index knee</td>
<td>4.8 (0.3)</td>
<td>4.7 (0.8)</td>
<td>.662</td>
<td>.510</td>
</tr>
<tr>
<td>Unaffected knee</td>
<td>4.9 (0.3)</td>
<td>4.7 (0.5)</td>
<td>2.46</td>
<td>.016*</td>
</tr>
<tr>
<td>Elbow</td>
<td>4.7 (9)</td>
<td>4.7 (1.1)</td>
<td>.211</td>
<td>.833</td>
</tr>
</tbody>
</table>
### Table A4.3: Comparison between CPT groups of effect size and Cohen's d for key study outcome measures.

<table>
<thead>
<tr>
<th></th>
<th>CPT &lt;15°C</th>
<th>CPT &gt;15°C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cohen's d</td>
<td>Effect size (r)</td>
</tr>
<tr>
<td>ADI</td>
<td>.291</td>
<td>.144</td>
</tr>
<tr>
<td>PainDETECT</td>
<td>1.31</td>
<td>.547</td>
</tr>
<tr>
<td>PPT mean</td>
<td>.877</td>
<td>.402</td>
</tr>
<tr>
<td>OA knee</td>
<td>1.03</td>
<td>.457</td>
</tr>
<tr>
<td>CPT mean</td>
<td>.439</td>
<td>.214</td>
</tr>
<tr>
<td>OA knee</td>
<td>.079</td>
<td>.040</td>
</tr>
<tr>
<td>WOMAC pain</td>
<td>1.20</td>
<td>.516</td>
</tr>
<tr>
<td>stiffness</td>
<td>1.59</td>
<td>.622</td>
</tr>
<tr>
<td>function</td>
<td>1.35</td>
<td>.558</td>
</tr>
<tr>
<td>ALF chair</td>
<td>.691</td>
<td>.327</td>
</tr>
<tr>
<td>walk</td>
<td>.852</td>
<td>.392</td>
</tr>
<tr>
<td>stairs</td>
<td>1.18</td>
<td>.510</td>
</tr>
</tbody>
</table>
Appendix 5

Further details of publications referenced in this Thesis


Appendix 5.5 Wright A, Kenner P, Moss P (2010): A proportion of normal subjects may respond abnormally to cold stimuli. IASP 13th World Congress on Pain, Montreal, Canada.

Appendix 5.6 Wright A, Ha L-B, Moss P, Benson HAE (2010): Development of a rapid acting topical delivery system for menthol as the basis for a new quantitative sensory test. Perspectives in Percutaneous Penetration 12th International Conference. La Grande Motte, France.

Appendix 5.7 Whitnall, J., Moss, P., & Wright, A. (Unpublished). Cold pain threshold testing is reliable in pain-free healthy adults.
Subjects with knee osteoarthritis exhibit widespread mechanical and cold, but not heat hyperalgesia

Penny Moss, Emma Knight, Anthony Wright

School of Physiotherapy, Curtin University of Technology, Perth, W.A., Australia

Background

Hyperalgesia to mechanical and cold stimuli is a recognized characteristic of neuropathic pain (1), fibromyalgia (2) and whiplash associated disorder (3).

There is scant data regarding osteoarthritis (OA) despite anecdotal evidence of adverse response to cold and movement. Although animal models of arthritis have demonstrated increased pain response to cold (4), few human studies have been conducted. More data is needed to confirm the presence of mechanical and thermal hyperalgesia in OA.

Method

Subjects: 23 subjects with knee OA and 23 age, gender and BMI-matched healthy controls were recruited voluntarily from the community (Table 1).

Inclusion/Exclusion Criteria: OA subjects fulfilled the ACR clinical diagnostic criteria for knee OA; all subjects - no history of additional joint problems, neurological deficit or chronic pain.

Design: cross-sectional; test site (Table 2) and modality order randomised.

Dependent Variables:

- Pressure pain threshold (PPT), measured using an electronic digital pressure algometer (Somedic, AB, Sweden), 1cm² probe applied at 40gPa/sec; subjects instructed to depress switch on change from pressure to pain; mean of 3 trials calculated for analysis.
- Cold pain threshold (CPT), measured using peltier thermode (Somedic, AB, Sweden) and method of limits, baseline 25°C, slope 1°C/sec, min 5°C; subjects instructed to depress switch on change from cold to pain; mean of 3 trials used for analysis.
- Heat pain threshold (HPT), measured as for CPT, baseline 32°C, slope 1°C/sec, max 50°C; mean of 3 trials used for analysis.

Ethics: All participants provided written informed consent and the study was approved by Curtin University Human Research Ethics Committee.

Data Analysis:

SPSS v(15) statistical package: α-level set at p>0.005 and β-level at 80%.

The following analyses were applied:

- Independent t-tests for between-group differences in PPT, CPT & HPT.
- Repeated measures ANOVA for within-subject differences in thresholds between sites.

Abstract number 3997

Results

Table 1: Gender, age and BMI group data

<table>
<thead>
<tr>
<th>Site</th>
<th>OA Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knee: OA</td>
<td>30.13</td>
<td>10.13</td>
</tr>
<tr>
<td>Age (yrs): mean ± SD (range)</td>
<td>68.5 ± 8.5 (55-82)</td>
<td>66.0 ± 11.5 (50-84)</td>
</tr>
<tr>
<td>BMI: mean ± SD</td>
<td>26.94 ± 4.51</td>
<td>25.61 ± 4.30</td>
</tr>
</tbody>
</table>

1. Subjects with OA showed widespread mechanical hyperalgesia

2. Subjects with OA demonstrated widespread cold hyperalgesia

Conclusions

This study provides clear evidence of widespread elevated mechanical and cold pain thresholds in subjects with symptomatic knee OA. There is no evidence of altered heat pain thresholds.

References


For further details, please email: P.Moss@curtin.edu.au

Figure 1: Mean PPT Values
Significant difference in PPT between OA and control subjects:
Knee: OAsubjects vs. control: t(25.3) = -2.37, p= 0.03
Ankle: OAsubjects vs. control: t(25.3) = -2.25, p= 0.04
Elbow: OAsubjects vs. control: t(25.3) = -2.15, p= 0.06
There was no significant difference in PPT values between sites:
OA subjects: F(2, 50) = 0.88, p= 0.43

Figure 2: Mean CPT Values
Significant difference in CPT between OA and control subjects:
Knee: OAsubjects vs. control: t(25.3) = 2.01, p= 0.05
Ankle: OAsubjects vs. control: t(25.3) = 2.03, p= 0.05
There was no significant difference in CPT values between sites:
OA subjects: F(2, 50) = 3.31, p= 0.047
Appendix 5.2

**Topical menthol evokes concentration-dependent cold, unpleasantness and pain responses in healthy subjects**

**Wright, A; Moss, P; Benson, H**

CHIRI, Curtin University of Technology, Perth, Australia.

**Aim of investigation**

Topical menthol has been proposed as an experimental model for cold hyperalgesia in humans. However there remain many questions regarding optimal application parameters and relationship to cold pain thresholds (CPT). This psychophysical study aimed to characterise the responses of healthy subjects to 3 different concentrations during a 30-min application and to investigate the relationship between response to menthol and CPT.

**Methods**

Thirty-two healthy adults (19-61 yrs) were voluntarily recruited. Using a blind randomised control design, four concentrations of menthol in 2mls solution were applied to the forearm via a 3x5cm gauze pad occluded by Tegaderm dressing, 4 test sessions were separated by at least 24 hrs. Every 60 secs during the 15-min application subjects rated cold, unpleasantness and pain intensity on labelled 10cm visual analogue scales (VAS). Area under each VAS-time curve was calculated, and repeated measures ANOVA used to analyse differences in intensity of cold, unpleasantness and pain for each concentration. CPT was tested using a peltier thermode (Somedic AB, Sweden) and method of limits, with baseline set at 32°C, slope 1°C/sec. The mean of 4 trials was calculated and Pearson’s Correlation Coefficient applied to assess associations between CPT and menthol responses (SPSS v15; α-level p>0.05).

**Results**

All 3 menthol concentrations evoked cold unpleasantness and pain sensations in a statistically significant concentration-dependent manner (cold: $F_{(2,62)}=11.32$, p>0.001; unpleasantness: $F_{(2,62)}=8.92$, p>0.001; pain: $F_{(2,62)}=4.21$, p=0.019). Sensation time-course also showed significant dose-dependency, with all mean times to onset for the lower concentration of menthol significantly later than for the higher concentrations. Mean time to peak pain also differed significantly between concentrations. Contrary to previous studies, even the lowest concentration of menthol evoked cold in 97% and pain in 25% of subjects. Significant correlations were found between CPT and responses to menthol. The ethanol control evoked no cold, unpleasantness or pain.
Conclusions

Topical menthol evoked a dose-specific response in normal subjects, with higher concentrations eliciting progressively greater cold, unpleasantness and pain. Response intensity also correlated with standard CPT testing, suggesting a clear association. The finding that low concentrations of menthol can activate the nociceptive system in normal subjects adds to the debate about the relative roles of TRPM8 and TRPA1 in cold hyperalgesia.
Appendix 5.3

**Subjects with lower limb osteoarthritis exhibit widespread cold hyperalgesia, but not heat hyperalgesia**

**A. Wright¹, E. Knight¹, S. McConnell¹, P. Moss¹**

¹CHIRI, Curtin University of Technology, Perth, Australia.

**Background:** Thermal hyperalgesia has long been acknowledged as characteristic of neuropathic pain. More recent studies have additionally suggested an association between chronic musculoskeletal conditions and cold hyperalgesia (1). Although animal models of arthritis have demonstrated thermal hyperalgesia, there have been few human studies. Further data would clarify the presence of hyperalgesia to cold and/or heat in osteoarthritis (OA) sufferers.

**Objectives:** To assess whether the cold pain (CPT) and heat pain thresholds (HPT) of subjects with hip or knee OA differ from those of matched healthy controls, and whether this response is localised or more widespread.

**Methods:** 23 subjects with knee OA and 18 subjects with hip OA, plus the same numbers of age, gender and BMI matched healthy controls were recruited voluntarily. Mean (SD) age: knee OA 68.4 (8.56) yrs, control 65.95 (11.12) yrs; hip OA 62.5 (9.04) yrs, control 58.9 (7.13) yrs. OA subjects fulfilled ACR clinical criteria for OA, with no history of additional joint involvement or chronic pain. A peltier thermode (Somedic, AB) was used to test CPT and HPT, applying standardised instructions and methodology. A practice and 3 trials were performed at each site for each measure in randomised order. CPT and HPT were tested at standardised sites: for knee subjects, at the knee, ipsilateral ankle and elbow; for hip subjects, at the affected hip, contralateral hip, ipsilateral knee and shoulder. Independent t-tests were applied to mean CPT and HPT values to analyse differences between groups at each site and globally (mean of sites). SPSS (v15) was used, α-level set at p>0.05. Ethical approval was granted by Curtin University HREC.

**Results:** The OA knee group showed significantly increased CPT at each site compared with controls (ankle: p>0.001; knee: p=0.030; elbow: p=0.035). Similarly, OA hip subjects exhibited CPT at temperatures higher than their healthy controls at each site.
However, whilst this difference was significant at both the affected hip (p = .034), and the contra-lateral hip (p = .040), it was not at the ipsilateral shoulder or knee (Table 1). In contrast, no difference in HPT was found between groups at any of the test sites for knee (means (SD): OA 48.09°C (1.75), healthy 47.94°C (1.94); p = .783) or hip: OA 46.47°C (3.64), healthy 47.61°C (2.14); p = .258.

Conclusions: Subjects with hip and knee OA exhibited significantly increased CPT, but no evidence of altered HPT, when compared with matched controls. This cold hyperalgesia was not restricted to the OA affected joint but in many subjects spread to the contra-lateral lower limb and even the upper limb, suggesting a more generalised change in nociceptive system sensitivity.

<table>
<thead>
<tr>
<th>Test sites</th>
<th>Mean (SD) CPT</th>
<th>% increase</th>
<th>t =</th>
<th>p =</th>
</tr>
</thead>
<tbody>
<tr>
<td>OA knee</td>
<td>8.1 (4)</td>
<td>11.2 (5.8)</td>
<td>37.5</td>
<td>2.08</td>
</tr>
<tr>
<td>IL ankle</td>
<td>7.1 (2.4)</td>
<td>12.0 (6.2)</td>
<td>68.7</td>
<td>2.04</td>
</tr>
<tr>
<td>IL elbow</td>
<td>8.9 (4.0)</td>
<td>12.5 (7.3)</td>
<td>39.6</td>
<td>3.31</td>
</tr>
<tr>
<td>GLOBAL</td>
<td>8.7 (4.0)</td>
<td>11.8 (5.4)</td>
<td>46.0</td>
<td>2.21</td>
</tr>
<tr>
<td>OA Hip</td>
<td>8.9 (5.7)</td>
<td>14.6 (9.2)</td>
<td>69.8</td>
<td>-2.21</td>
</tr>
<tr>
<td>CL hip</td>
<td>5.8 (4.1)</td>
<td>10.7 (8.7)</td>
<td>84.5</td>
<td>-2.14</td>
</tr>
<tr>
<td>IL knee</td>
<td>9.6 (6.2)</td>
<td>13.1 (8.0)</td>
<td>36.5</td>
<td>-1.46</td>
</tr>
<tr>
<td>IL shoulder</td>
<td>9.2 (7.5)</td>
<td>12.0 (9.7)</td>
<td>26.1</td>
<td>-.964</td>
</tr>
<tr>
<td>GLOBAL</td>
<td>8.8 (5.2)</td>
<td>13.6 (7.8)</td>
<td>47.8</td>
<td>-1.98</td>
</tr>
</tbody>
</table>

Table A5.1: Differences in cold pain threshold between healthy and OA subjects
Appendix 5.4

**Hip osteoarthritis subjects with cold hyperalgesia report dysesthetic and paradoxical sensations.**

**P.Moss¹, C.Pickard¹, S.McConnell¹, A.Wright¹**

¹School of Physiotherapy, Curtin University of Technology, Perth, Australia.

**Background:** Cold hyperalgesia with paradoxical dysesthesia is considered characteristic of neuropathic pain. Recent studies have also reported sensory hyperalgesia with neuropathic qualities (“burning”, “stinging”) in subjects with osteoarthritis (OA) (1). This study aimed to explore the quality of sensation experienced at cold pain threshold (CPT) by subjects with hip OA.

**Objectives:** To characterise the sensory experience at CPT for subjects with hip OA. To evaluate i) differences in descriptors selected by those with higher CPT; ii) associations between CPT and descriptor index scores.

**Methods:** 18 subjects with mild to moderate hip OA were recruited voluntarily: mean age 62.5 (SD 9.04) years. Subjects fulfilled the ACR clinical criteria for hip OA, with no history of additional joint involvement or chronic pain. A peltier thermode (Somedic, AB) was used to provide a cold stimulus to the CPT of each subject applying standardised instructions and methodology. Following 1 practice and 3 trials at an anterior hip site, subjects were asked to select from the McGill Pain Questionnaire (2) whichever words best described the sensation at CPT. Post hoc, subjects were grouped according to CPT < or > 12°C. Chosen words were categorised according to severity (0-5) using the recently developed Algoteect Descriptor Index (ADI). The established indices Pain Rating Index (PRI) and Number of Words Chosen (NWC) were also calculated. Independent t-tests analysed differences between CPT groups. Pearson’s correlation coefficient assessed associations between CPT and descriptor indices. SPSS (v17) was used, α-level p>0.05. Ethical approval granted by Curtin University HREC.

**Results:** 56% of subjects with hip OA exhibited a CPT of greater than 12°C (mean 19.48°C, SD 4.97°C). Whilst all subjects with more normal CPT (<12°C) described the sensation as predominantly cool or cold, those with higher CPT chose a range of more unpleasant words, particularly prickling, stinging and burning, with only 40% choosing cold thermal words. Statistical analysis revealed that subjects in the hyperalgesic group
scored significantly higher on the ADI (p = .005). Although PRI scores were 78% higher for the >12°C group, high variance prevented statistical significance (p = .098). There was no significant difference in NWC (p = .188). Pearson’s coefficients showed significant correlations between ADI and PRI (r = .615, p = .007) and between CPT and both descriptor indices (ADI r = .710, p = .001; PRI r = .477, p = .045). There were no significant correlations between NWC and CPT.

**Conclusions:** 56% of subjects with mild to moderate hip OA exhibited cold hyperalgesia, strongly associated with higher values on both new and standard descriptor indices. These subjects were likely to describe this threshold sensation as prickling, stinging or burning, words that are more often associated with neuropathic rather than musculoskeletal pain. Larger studies are needed to substantiate these findings.

1. Hochman et al., IASP Glasgow, 2008

<table>
<thead>
<tr>
<th>CPT</th>
<th>PRI Mean (SD)</th>
<th>PRI t =</th>
<th>PRI p =</th>
<th>NWC Mean (SD)</th>
<th>NWC t =</th>
<th>NWC p =</th>
<th>ADI (0-5) Mean (SD)</th>
<th>ADI (0-5) t =</th>
<th>ADI (0-5) p =</th>
<th>CPT °C Mean (SD)</th>
<th>CPT °C t =</th>
<th>CPT °C p =</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 12°C</td>
<td>1.63 (.74)</td>
<td>-1.80</td>
<td>.098</td>
<td>1.25 (.46)</td>
<td>-1.38</td>
<td>.188</td>
<td>1.44 (.73)</td>
<td>-2.88</td>
<td>.011</td>
<td>6.29 (2.23)</td>
<td>629</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>&gt; 12°C</td>
<td>2.90 (2.1)</td>
<td>1.70 (1.82)</td>
<td>2.77 (1.13)</td>
<td>19.48 (4.97)</td>
<td>194.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A5.2: Differences between subjects with CPT < or >12°C
A proportion of subjects may respond abnormally to cold stimuli

Kenner P, Moss P, Wright T
Curtin Health Innovation Research Institute, School of Physiotherapy, Curtin University, Perth, W.A., Australia

Background
Cold hyperalgesia, as defined by increased cold pain threshold (CPT), has long been associated with neuropathy and more recently with chronic musculoskeletal pain (1). Since there is no clear agreement about what characterizes an abnormal response to cold, further normative data is needed. This study aimed to explore cold pain threshold and descriptive data for healthy subjects, comparing age and gender groupings.

Key Study Questions:
1. What is the mean response of healthy adults to a self-regulated transitory cold stimulus (CPT) and descriptor choice?
2. Is there a difference between upper limb sites?
3. Is there a difference between males and females?
4. Does cold response vary according to age?

Method
• 66 healthy, pain-free subjects (31 male, 37 female) were recruited equally across 3 age bands (18-29, 30-49, 50+ yrs).
• A peltier thermode (Somedic AB, Sweden) was used: 3x3cm probe, baseline 32°C, ramp 2°C, min 5°C. Method of limits was applied, using standardized instructions:
  • CPT was measured at 4 standardised upper limb sites: deltoid bilaterally, ipsilateral elbow and wrist – Fig 1a) in random order;
  • one practice was followed by 3 trials.

• Immediately following CPT, subjects chose descriptors from a McGill Pain Questionnaire list. Word choices were categorized post hoc according to severity using:
  a) The recently developed Algotect Descriptor Index (ADI);
  b) The MPQ Pain Rating Index (PRI);
  c) The MPQ Number of Words Chosen (NWC).
  For all indices, low score = milder response.

Ethical approval was granted by Curtin University HREC.

Data Analysis:
• SPSS (v17) was used, p<0.05;
• Gender differences were analysed using independent t-tests;
• Differences between age bands was analysed using one-way ANOVA;
• Pearson’s correlation coefficients assessed associations between CPT and descriptor indices (ADI, PRI, NWC).

Quality of sensation at CPT
95% subjects chose at least 1 cold-thermal word to describe the CPT sensation. 40% also chose a paradoxical thermal word such as ‘hot’ or ‘burning’. Two-thirds of those who reported paradoxical heat also described the sensation as unpleasantly ‘prickling’ or ‘tingling’. Only 10% of this group selected the milder ‘tingling’.

Gender differences in word choice: There were only minor differences, with slightly more females choosing ‘burning’ or ‘hot’, and more males choosing ‘cool’ or ‘cold’.

No statistically significant difference between genders was found when ADI, PRI and NWC scores were compared (p= .158, p= .804, p= .973).

Differences between age bands was analysed using one-way ANOVA; there was no significant difference in CPT between males and females.

CPT and descriptors: There was a moderate positive correlation between CPT and ADI (r= .499, p< .001) and PRI (r= .294, p= .0.15) descriptor index scores. When subjects were grouped according to CPT < or > 12°C (a possible ‘normal cut-off’, differences in word choice emerged (Fig 7). This was reflected in significantly different ADI scores between groups (p > .001).

Conclusions
The majority of healthy subjects experienced CPT at 12°C or lower, describing this sensation as predominantly ‘cool’, ‘cold’, ‘dull’ or ‘numb’. However, up to 1/3 may exhibit CPT of above 12°C. This group select paradoxical and unpleasant heat or dysesthetic words. No significant difference in threshold or word choice was found according to gender or age.

References:
1. Sterling et al., 2006

Fig 1a: Upper limb test sites
Fig 1b: Thermode test set up

Mean CPT values:
<table>
<thead>
<tr>
<th>Site</th>
<th>18-29 yrs</th>
<th>30-49 yrs</th>
<th>50+ yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deltoid (F)</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
<tr>
<td>Deltoid (M)</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
<tr>
<td>Elbow (F)</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
<tr>
<td>Elbow (M)</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
<tr>
<td>Wrist (F)</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
<tr>
<td>Wrist (M)</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
<tr>
<td>D/C/L</td>
<td>4-50.0</td>
<td>4-50.0</td>
<td>4-50.0</td>
</tr>
</tbody>
</table>

Fig 4: Mean CPT at each site, according to gender

Fig 5: Mean CPT at each site, according to gender

Fig 6: Mean CPT at each site, according to age

Fig 7: Descriptor index scores, compared by age

Fig 8: Word choices according to CPT grouping

Fig 9: Most frequently chosen descriptors.

Fig 10: Distribution of overall CPT (mean of all sites)

Fig 11: Distribution of overall CPT (mean of all sites)

Fig 12: Distribution of overall CPT (mean of all sites)
Appendix 5.6

Development Of A Rapid Acting Topical Delivery System For Menthol As The Basis For A New Quantitative Sensory Test

A Wright1,2, L-B Ha2, P Moss1,2, H Benson1,2

1School of Physiotherapy, 2 School of Pharmacy, 3Curtin Health Innovation Research Institute, Curtin University, Perth, WA.

Objective To develop and characterise a rapid release topical menthol delivery system for application to the skin. To evaluate the sensory response to different menthol concentrations. Menthol could be used to provide a standardised sensory stimulus for quantitative sensory testing in patients with chronic pain.

Methods Menthol formulations consisting of organic solvent based solutions, gels and sprays were prepared. The release of menthol from each formulation was determined using a Permegear® ILC14 Automated System with flow-through cells. Experimental parameters to determine diffusion through a semi-permeable cellulose membrane were: 35°C, flow rate at 55μL/min and 50%v/v aqueous ethanol as receptor solution: n=6 for each formulation. Receptor solution samples were collected at 10 minute intervals for 60 minutes and menthol content analysed by gas chromatography. In a separate study, menthol gel formulations (Gel A and Gel B) were applied to the volar surface of the forearm in healthy human volunteers (n=27). Subjects rated the intensity of their sensory responses (Cold, Heat, Unpleasantness and Pain) to the menthol stimulus using 100mm electronic visual analogue scales (VAS) at 2 min intervals over a 20 min period.

Results Menthol release from the formulations ranged from 158.9 ± 09.5 mg to 454.7 ± 29.4 mg menthol 0-60 min from the Gel B menthol formulations. Spray and gel formulations containing complex solvent systems released significantly more menthol than simple aqueous ethanol solutions. In human volunteers, both gel formulations produced sensory responses. Cold sensation was reported by almost all subjects for both formulations, although no difference was found between formulations in cold intensity ratings (p=.386). In contrast, differences were noted for heat (p=.050), unpleasantness (p=.010) and pain (p=.012) between concentrations, with Gel B producing significantly higher intensity ratings.
Discussion Rapid menthol release topical formulations were developed and assessed in vitro. Formulation solvent vehicles influence menthol release with significantly enhanced menthol release achieved. The release differences are likely to be related to thermodynamic activity and solubility differences of menthol in the vehicles. Application of menthol to the skin using a rapid release formulation produced a range of sensory responses including heat and pain as well as the normal cold sensation. Response intensity is concentration dependent. Topical menthol application may therefore provide the basis for a simple quantitative sensory test for cold allodynia that would have clinical value in assessing patients with chronic pain.

Acknowledgements This study was funded by a Curtin University of Technology Internal Research Grant. Formulation development was assisted by MedPharm Ltd.
Appendix 5.7

Cold Pain Threshold Testing Is Reliable In Pain-Free Healthy Adults.

J. Whitnell, P. Moss, & A. Wright (Unpublished)

School of Physiotherapy, Curtin Health Innovation Research Institute, Curtin University, Perth, WA.

Background: Increased sensitivity to cold occurs in many chronic pain conditions but there is limited evidence as to the reliability of cold pain threshold (CPT) or the effect of gender or psychological factors on CPT in healthy adults. This study aimed to determine the reliability of CPT across three separate sessions and the effect of gender, anxiety and pain catastrophising on baseline CPT in healthy subjects.

Methods: CPT testing using the standard Method of Limits was performed on 45 healthy adults (M:F 20:25) on three separate test sessions, each at least one day apart, at four test sites: thenar eminence; volar forearm; tibialis anterior; plantar foot. Descriptor choice from the Short-Form McGill Pain Questionnaire-Part 2 and VAS pain intensity rating at CPT were recorded each session. Anxiety (State-Trait Anxiety Inventory) and pain catastrophising (Pain Catastrophising Scale) levels were recorded before CPT testing.

Results: CPT, descriptors and VAS at CPT were found to be highly reliable over three test sessions (ICC ranges r= .87-.93; r=.85-.89; r= .94-.97). Female CPT was more reliable and showed a significantly higher sensitivity to cold at most sites. There was no significant correlation between anxiety measures and CPT (r= -.29-.09, p= 0.053-0.779). Similarly, there was no correlation between pain catastrophising measures and the thenar eminence, volar forearm or plantar foot.

Conclusion: CPT testing alongside descriptor choice and VAS pain intensity are relatively reliable tools for testing response to cold in healthy adults. Future studies are needed to assess reliability in populations with pathology.