CONTESTED CREDIBILITY OF FIBROMYALGIA WITHIN MEDICAL, LEGAL,
AND INSURANCE CONTEXTS

by

JUDY ANN LE PAGE

B.A. (Hons.), University of Manitoba, 1983
M.A., University of Manitoba, 1985

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

INTERDISCIPLINARY STUDIES

THE UNIVERSITY OF BRITISH COLUMBIA

July 2005

© Judy Ann Le Page, 2005
Credibility and Fibromyalgia

Abstract

Fibromyalgia (FM) is a controversial diagnosis that has been associated with considerable personal and occupational disability. The difficulty in determining the credibility of reported pain and disability in FM has perplexed the various disciplines involved in this process. The purpose of this dissertation was to take a multidisciplinary approach and examine the complexities of FM as faced by patients, the medical and legal system, and the insurance industry. Two independent, yet related, studies comprise this dissertation, each examining aspects of the various perspectives involved in disability claims for FM. Study 1 was more exploratory in nature and involved a systematic review of every trial-by-judge litigated FM claim in Canada (N=194) up to 2003. The cases were examined in relation to demographic, insurance, and credibility factors. The demographic factors revealed a gender disparity of women (84%) plaintiffs, and a disproportionate number of FM cases litigated in British Columbia (61%). Insurance related factors indicated that although surveillance information played a role in disability determinations for FM, the credibility of that information was central to the amount of award granted. Plaintiff credibility played a similar role, indicating that plaintiffs perceived as more credible were typically granted greater awards. An examination of medical expert credibility revealed that judges appear to perceive experts as more credible overall than plaintiffs, regardless of the expert’s familiarity level with the plaintiff. The purpose of Study 2 was to examine the phenomenology of exaggeration and the deliberate portrayal of excessive disability in patients with FM. Fifty-four patients with FM completed a variety of psychological, physical, functional, and symptom validity measures twice. In one session, patients were instructed to complete the tests and measures in a manner that was “normal” for them, while patients were instructed to “exaggerate” pain and disability associated with FM in the other session. There were medium to large differences associated with exaggeration on some of the self-report measures of physical, functional, and psychological status. The subjects did not, however, exaggerate on the symptom validity tests. Although multiple cutoff scores exceeded single test scores in the identification of exaggerators in most cases, the false negative rate remained moderate to high. These findings suggest that patients with FM can exaggerate their disability on self-report measures in a plausible manner.
# TABLE OF CONTENTS

Abstract .......................................................................................................................... ii
Table of Contents ........................................................................................................ iii
List of Tables ................................................................................................................ v
List of Figures ............................................................................................................... vii
Acknowledgements ..................................................................................................... viii
Introduction .................................................................................................................. 1

## Medical, Psychological, and Epidemiological Issues

- Diagnostic Criteria and Accuracy ........................................................................... 4
- Prevalence & Gender ................................................................................................. 7
- Biomedical Factors ................................................................................................... 9
- Psychological Factors ............................................................................................... 18
- Summary ................................................................................................................... 25

## Study 1: FM in the Canadian Courts

- Insurance and Disability Claims Issues ................................................................. 29
- The Insurance Industry .......................................................................................... 30
- Disability Insurance Law ....................................................................................... 34
- Proof of Disability ................................................................................................... 35
- Expert Evidence ...................................................................................................... 38

## Credibility and Fibromyalgia

- Mitigation, Contributory Negligence, Aggravation, and Punitive Damages ......... 40
- Material Misrepresentation and Fraud ................................................................. 41
- Summary ................................................................................................................... 43
- Hypotheses ............................................................................................................... 45
- Method ....................................................................................................................... 47
- Reliability ................................................................................................................ 48
- Statistical Analyses ................................................................................................. 49
- Preliminary Considerations Regarding Statistical Analyses ................................ 49
- Results ....................................................................................................................... 50
- Sample Characteristics ........................................................................................... 50
# Table of Contents

Gender Related Factors ......................................................... 55  
Insurance Factors .......................................................... 56  
Credibility Factors ........................................................... 59  
Discussion ................................................................. 65  
Limitations and Considerations for Future Research ............... 70  
Study 2: FM and Exaggerated Disability .................................. 72  
Patient Perspectives ....................................................... 73  
Compensation & Prevalence Rates ........................................ 75  
Compensation Status & Chronic Pain .................................... 76  
Symptom Exaggeration & Chronic Pain .................................. 78  
Effort Testing ................................................................. 84  
Methodological Considerations in Effort Testing ...................... 86  
Hypotheses ................................................................. 88  
Method ................................................................. 89  
Participants ............................................................... 89  
Diagnostic Assessment ..................................................... 89  
Measures ................................................................. 90  
Procedures ............................................................... 96  
Statistical Analyses .......................................................... 99  
Results ................................................................. 100  
Evaluation of Psychological Test Performance ......................... 100  
Evaluation of Personality Assessment Inventory Performance ...... 101  
Evaluation of Effort Test Performance .................................. 102  
Examination of Frequency Distributions ................................ 105  
Post-Experimental Questionnaire ........................................... 112  
Health Insurance Inventory ................................................ 113  
Discussion ............................................................... 115  
Limitations and Considerations for Future Research .................. 139  
General Discussion .......................................................... 141  
References ............................................................... 145  
Appendix A: Study 1: Variable Coding Sheet ......................... 185  
Appendix B: Post-Experiment Interview ................................. 187  
Appendix C: Health Insurance Inventory ................................. 188
List of Tables

Table 1. Awards & damages granted for total sample .......................................................... 51
Table 2. Awards & damages granted by geographic location: western Canada ...................... 52
Table 3. Awards & damages granted by geographic location: eastern Canada ....................... 52
Table 4. Summary of awards compared across three provinces ........................................... 54
Table 5. Number of cases litigated over time by geographic location .................................. 55
Table 6. Summary of credibility ratings and total awards granted by gender ......................... 55
Table 7. Total award granted relative to the presence of liability issues ............................... 57
Table 8. Specific awards and damages granted relative to the presence of liability issues ....... 58
Table 9. Amount of award granted relative to presence of surveillance information ............... 58
Table 10. Amount of award granted relative to plaintiff and surveillance credibility ................. 59
Table 11. Plaintiff credibility relative to the presence of liability issues ............................... 61
Table 12. Plaintiff credibility relative to expert credibility ratings ...................................... 61
Table 13. Percentage of plaintiffs and experts rated as credible ......................................... 62
Table 14. Expert credibility rating relative to treating clinician versus forensic expert .......... 63
Table 15. Expert credibility rating relative to plaintiff versus defense expert ......................... 64
Table 16. Psychological test scores by experimental condition ........................................... 100
Table 17. Multidimensional Pain Inventory test scores by experimental condition ................. 101
Table 18. Personality Assessment Inventory scale and index scores by experimental condition .......................................................... 102
Table 19. Effort test scores by experimental condition for all patients (N=54) ....................... 103
Table 20. Effort test scores by experimental condition only for exaggerators (n=12) ............... 105
Table 21. Cutoffs, accuracy, and false positive rates for psychological test scores ................ 106
Table 22. Cumulative frequency (%) of score combinations for the BDI, FIQ, and ODI ........ 107
Table 23. Cumulative frequency (%) of score combinations for the BDI, FIQ, ODI, and MPI-Pain Severity scale .......................................................... 107
Table 24. Cumulative frequency (%) of score combinations for the BDI, FIQ, ODI, PSQI, and MPI-Pain Severity, Interference, and Affective Distress scales ......................... 108
Table 25. Cumulative frequency (%) for score combinations for the MPI-Pain Severity, Interference, and Affective Distress scales .......................................................... 108
Table 26. Cutoffs, accuracy, and false positive rates for Personality Assessment Inventory scales and indexes ................................................................. 109
Table 27. Cumulative frequency (%) of score combinations for the NIM, MAL, and RDF ...... 110
Table 28. Cumulative frequency (%) of score combinations for the NIM and MAL ............. 111
Table 29. Cumulative frequency (%) of score combinations for the Depression, Anxiety, Anxiety-Related, Somatic Complaints, Schizophrenia, and Borderline Features .......... 111
Table 30. Cumulative frequency (%) of score combinations for the Depression, Anxiety, Anxiety-Related, and Somatic Complaints scales ................................................................. 112
Table 31. Cumulative frequency (%) of score combinations for the NIM, MAL, and selected clinical scales ................................................................................ 112
Table 32. Patient responses to insurance policy coverage for unreported or exaggerated health information ................................................................. 114
Table 33. Beck Depression Inventory-II scores from the literature ........................................ 116
Table 34. Fibromyalgia Impact Questionnaire scores from the literature ......................... 117
Table 35. Oswestry Disability Index 2.0 scores from the literature ....................................... 118
Table 36. Pittsburgh Sleep Quality Index scores from the literature .................................... 119
Table 37. Multidimensional Pain Inventory – 1 test scores from the literature ................... 121
Table 38. Multidimensional Pain Inventory – 1 test scores for patients with FM and comorbid depression and anxiety ................................................................................. 121
Table 39. Personality Assessment Inventory scores from the literature & Study 2 ............. 125
Table 40. Personality Assessment Inventory scores from the literature ............................... 126
Table 41. Personality Assessment Inventory scores for Feigning Mental Disorders .......... 126
Table 42. Classification rates (%) from test scores of assessed malingering in pain patient samples from the literature ................................................................. 129
Table 43. Victoria Symptom Validity Test scores from the literature .................................. 131
Table 44. Test of Memory Malingering scores from the literature ....................................... 132
Table 45. 21 Item Test scores for Feigning Mental Disorders ............................................. 134
List of Figures

Figure 1. Fibromyalgia Impact Questionnaire scores in relation to the literature. ..................117
Figure 2. Oswestry Disability Index 2.0 scores in relation to the literature. ..........................118
Figure 3. Pittsburgh Sleep Quality Index scores from the literature ....................................119
Figure 4. Pain Severity scores from the literature .................................................................122
Credibility and Fibromyalgia

Acknowledgements

There are a number of individuals I wish to thank and acknowledge for their assistance and support during the completion of this dissertation and my graduate studies.

Thanks are extended to Dr. Grant Iverson, my research supervisor, for his guidance and expertise in this complex project. I also thank other members of my research committee. Dr. Barry Koehler, Dr. Kenneth Craig, Mr. Peter Collins, and Dr. Peter Darke each brought their perspectives and disciplines to my research, and each contributed to my breadth of knowledge, and greatly enhanced the quality of my dissertation. Thanks are also extended to my university examiners, Drs. William Koch and Izabela Schultz, and my external examiner, Dr. Ken Prkachin, for their helpful comments.

I am indebted to the student research assistants who spent many hours assisting me with various aspects of this dissertation, including Sundeep Gill, Jessica Broderick-Lockhardt, Meagan Hasek-Watt, and especially to Christian and Jordan Maile for their endless assistance. I thank Drs. Barry Koehler, Kam Shojania and Maziar Badii for their medical expertise and time in diagnosing patients for Study 2. Thanks also go to Jennifer Bernardo for her technical support and assistance. Thanks are extended to Mary Koehler, Cindy Rosenbaum, and Jan Wilson for coordinating office space and schedules, and the Arthritis Research Centre of Canada for providing office space for Study 2. Thanks also go to the Law Library staff at UBC, for facilitating access to the legal database. I am grateful to the patients with Fibromyalgia who participated in this research.

In addition, I would like to thank my mother, Helen Bainski, for her continual love and support, and my brother, Rob Zaparniuk, for his helpful suggestions and hours of discussion on the legal aspects of my research. I also thank several dear friends who have supported me through this journey, especially Maureen Whittal, Clare Phillips, Wendy Dawson, Lily Rosenberg, and Ken Le Page.

On a more personal note, I am forever grateful to my husband, Tony Le Page, whose love, friendship, and enthusiasm was central in maintaining my persistence. I am also most grateful to my very best buddy Max, who kept me sane and always reminded me of the important things in life. I would like to dedicate this dissertation to Tony and Max, for I could not have achieved this without them.
Introduction

Fibromyalgia (FM) is a confusing and controversial diagnosis. This syndrome is characterized by widespread pain and by tenderness at specific anatomical sites, and patients frequently complain of diverse physical and psychological problems. The cause of this syndrome is unknown, and the course of the condition is difficult to predict. The reality is that many patients with FM become progressively worse, leading to disability. Given that there is no obvious cause and no proven cure for the syndrome, it is not surprising that FM is a contentious diagnosis from a medical perspective, as well as in the disability insurance industry and in civil litigation.

From a medical perspective, FM has been associated with controversy since receiving a diagnostic classification and criteria in 1990. Even though the diagnostic criteria and label continues to be applied, there has been much speculation and debate in the medical community regarding the utility of some of the criteria in evaluating FM. Numerous investigations have been conducted from a variety of perspectives in the medical research, but all have been inconclusive thus far. Similarly, the psychological literature has offered various perspectives, and the findings have not consistently supported one particular view.

From an insurance industry perspective, FM is one of the most problematic conditions due mainly to the substantial impact it can have on an individual’s physical functioning, disability level, and quality of life. Consideration must also be given to the increase in healthcare costs, disability insurance claims, and resulting economic burden that has been associated with some of the chronic pain disorders for which there is no observable damage (e.g., whiplash, FM; Malleson, 2002). Fraud and misrepresentation are also potential sources of increased cost to the insurer, especially in situations where objective evidence is not available.

In chronic pain patients, it is relatively well established that identifiable, underlying pathophysiology is frequently absent. Only a limited association exists between confirmed pathology and self-reported pain (e.g., cancer and back pain; Deyo, 1986; Syrjala & Chapko, 1995), and physical pathology has not been predictive of disability (e.g., rheumatoid arthritis; Hagglund, Haley, Reveille, & Alarcon, 1989). Robinson (2001) summarized the physical pathology and disability issue by stating that even when patients have objective findings, those findings seldom explain the extent of the incapacitation reported by the patient.
Consequently, in the absence of observable damage, chronic pain conditions such as FM present a special challenge to the court system because they do not appear to exist, and they challenge the belief systems of people unfamiliar with such conditions. Furthermore, determining the credibility of pain complaints is a complex process in a system where, as Oatley (1997) noted, there is resistance to believing in something which you cannot see and for which there is no obvious evidence. Even though the current state of the law indicates that no particular type of evidence is required in order to prove disability, the best evidence of disability, from either the insurer’s or the insured’s perspective, remains objective evidence. The courts have not reached a consensus regarding the validity of FM, and as such, plaintiff credibility is an increasingly important issue in the context of disability claims from legal and insurance industry perspectives (Ericson & Doyle, 2004; Wolfe & Potter, 1996).

The difficulty in determining the credibility of reported pain and disability level continues to perplex the medical, legal, and insurance industry, even though it is well known that an individual can experience pain in the absence of an injury or illness (e.g., headache, toothache, and back pain). The research has generally maintained a narrow focus on individual theoretical models, with no attempt at integrating those models. Some researchers support the view that FM is multifaceted, with an interplay of biological and psychological factors. The patient, the doctor, the medical profession, the scientific literature, the insurance carrier, and the defendants in litigation must be considered. Each has a unique perspective. Each has expectations and responsibilities. Research in multiple domains is needed to fill in the missing pieces of this profoundly puzzling and controversial disorder.

The purpose of this dissertation was to examine the complexities faced by patients with FM, the medical system, the legal system, and the insurance industry. A multidisciplinary approach was taken to expand our understanding of the diverse perspectives that are part of disability determinations in patients with FM, and to examine how each perspective may interface with the other. Two independent, yet related, studies comprise this dissertation, each attempting to capture different aspects of the various perspectives involved in disability claims for FM. The first study focuses on the legal and insurance industry perspectives, while the second study focuses on the psychological and medical perspectives.
To appreciate the challenges associated with establishing disability due to FM, one must have a basic understanding of the components that comprise the system that must be navigated by the patient. The first section of this dissertation provides an overview of the components in this system, including an introduction to FM, epidemiological factors, diagnostic criteria, hypothesized pathophysiology, and psychological correlates. The second section contains a description of Study 1, including an overview of issues related to the insurance industry and disability insurance law, the purpose, hypotheses, method, results, and discussion. The third section contains a description of Study 2, including an overview of the research related to the exaggeration of medical conditions and effort testing, the purpose, hypotheses, method, results, and discussion. The fourth and final section contains an overall summary and integration of the results from both studies, and directions for future research.
Medical, Psychological, & Epidemiological Issues

FM is a controversial disorder of uncertain etiology, characterized by diffuse musculoskeletal pain and tenderness at specific anatomical sites. Patients with FM typically report a range of functional limitations (e.g., personal and home care difficulties), muscle stiffness, fatigue, irritable bowel syndrome, headaches, and psychological dysfunction, including sleep disturbance, depression, anxiety, and stress (Baumstark & Buckelew, 1992; Smythe, 1989; Uveges et al., 1990; Wolfe et al., 1990). The tender sites were recognized in the early 20th century and referred to then as “fibrositis”, as coined by Sir William Gower in 1904 (Osler, 1926). The natural course of FM is neither progressive nor fatal. Despite the absence of any definitive pathology, FM has been reported to significantly compromise physical functioning and quality of life, and be among the most disabling chronic disorders (Burckhardt, Clark, & Bennett, 1993; Hawley & Wolfe, 1991). For example, Wolfe et al. (1997a) followed FM patients for up to seven years, and found that although functional disability worsened slightly and health satisfaction improved slightly, measures of pain, global severity, fatigue, sleep disturbance, anxiety, depression, and health status were markedly unchanged. Due to the absence of observable tissue pathology or definitive biological markers in FM, the diagnosis has relied on physical signs and patients’ reports of symptoms.

Diagnostic Criteria & Accuracy

In 1990, a multicentre study conducted in the United States and Canada demonstrated that two criteria best identified FM accurately (Wolfe et al., 1990). Those criteria include: (a) the presence of widespread pain of at least three months duration, and (b) pain and tenderness reports upon palpation at 11 or more of 18 specific areas, referred to as ‘tender points’. The American College of Rheumatology (ACR) proposed a classification system comprised of those two criteria for the identification and diagnosis of FM. The specified tender points have been found to distinguish fibromyalgia from other soft tissue rheumatic disorders (McCain & Scudds, 1988).

There has been some disagreement in the literature regarding the diagnostic specificity of the FM tender points, and the association between tender points, pain complaints, and disability. Jacobs et al. (1996) found no association between self-reported pain and pain rating scores given on tender points, whereas, Croft et al. (1996) found an
association between tender points and reports of painful body segments. Wolfe (1998) found that even though 63% of his FM population reported pain in non ‘tender point’ regions called ‘control points’, the ‘tender points’ continued to be significantly better at differentiating FM patients from those without FM. White, Harth, Speechley, and Østbye (2000) found that adults meeting the ACR criteria for FM had a mean total tender point count more than 2-fold greater than pain controls with widespread musculoskeletal pain. They also found that all of the 18 tender points were more likely to be tender for FM cases than for pain controls.

One study compared pain ratings for each of the ACR-designated tender and control point areas in patients with FM to patients with other pain syndromes and to healthy controls (Okifuji, Turk, Sinclair, Starz, & Marcus, 1997). An evaluation of pain ratings across groups revealed pressure pain severity was significantly different between tender and control point areas. Patients with FM reported significantly higher levels of pain for both tender and control point areas, compared to patients with chronic back pain, chronic headaches, rheumatoid arthritis, and healthy controls. Similarly, patients from the other pain syndrome groups reported significantly higher levels of pain for both tender and control point areas, compared to healthy controls. Wolfe (1998) indicated that it is not rare to encounter a patient who finds every palpated site painful, and suggested that the increased responses might reflect exaggeration, or anxiety, anticipated pain, or a very low pain threshold. Aronoff (1998) suggested that deliberate exaggeration or obvious secondary gain should not automatically be concluded from excessive pain behaviour, but rather that it is part of a conditioned process of illness behaviour.

Some writers have expressed serious concern about the diagnosis of FM; Norton Hadler has expressed the most extreme position in this regard. Hadler (1999) described FM as a syndrome of “being out-of-sorts”, whereby those afflicted by the physical symptoms seek medical guidance rather than devise coping schemes. Burnum (1978) described the labeling process of such syndromes as FM as one that results in individuals joining the ranks of the “worried sick”. Hadler does not consider the FM diagnostic criteria, as well as criteria for chronic fatigue, myofascial pain, and irritable bowel syndrome, as possessing any significant utility, and regarded them as a form of illness behaviour magnified by the medicalization process. In addition to providing support for Hadler’s views, Hazemeijer and Rasker (2003) question the existence of FM through a philosophical analysis of the
diagnostic process. Hadler (1996) suggested that the very act of labeling a patient's pain state legitimizes unproven pathophysiological theories in ways that can be counterproductive, and that effective management is replaced by an adaptation of the sick role. Aronoff (1998) concurred with Hadler that the diagnostic labeling process may encourage illness behaviour and contribute to dysfunction in vulnerable individuals for conditions such as FM and myofascial pain syndrome. However, Aronoff further indicated that these diagnoses are appropriate for some patients, and that it is the responsibility of the physician to encourage patients not to get "trapped" in a counterproductive role. Writers at the opposite end of the continuum have expressed the view that FM does exist, and support the utility in diagnosing it (e.g., Goldenberg, 2004; Nielson & Harth, 2004; White, 2004).

In a prospective study examining the effects of assigning the FM label, researchers compared a group of patients with FM who had previously been diagnosed with FM, to a group who had not been previously diagnosed with FM (White, Nielson, Harth, Ostbye, & Speechley, 2002). They reported that although patients with a prior diagnosis of FM tended to be clinically worse than the newly diagnosed group, the newly diagnosed group did not worsen over time (36 month follow-up). Although the newly diagnosed patients appeared to become somewhat less active over time, there was a significant improvement in satisfaction with health, and they reported fewer symptoms and major symptoms lessened over time. The authors concluded that the FM label did not have a meaningfully adverse affect on clinical outcome over time, and that the FM label may have allowed for more appropriate medical management. A review of studies examining FM prevalence rates similarly suggests that there is no evidence that the FM label causes illness behaviour, or increases dependence on the medical system (White & Harth, 2001).

Given that FM is a compensable condition in both tort litigation and under disability insurance contracts, there can be considerable financial incentive to exaggerate or even fabricate disability. As such, there has been interest in determining whether physicians can differentiate real from fabricated FM. Khostanteen, Tunks, Goldsmith, and Ennis (2000) assessed physicians' ability to differentiate between subjects with FM and participants simulating FM. Although there was an approximate accuracy rate of 70% in correctly classifying patients with FM and simulators, simulators were misclassified as having FM about one third of the time, while patients with FM were misclassified as simulators almost
as frequently. Similarly, Smythe and colleagues reported a 60% accuracy rate for detecting simulators (thus, a 40% false positive rate) and an 86% accuracy rate for patients with FM (thus, a 14% false negative rate; Smythe Gladman, Mader, Peloso, & Abu-Shakra, 1997). These studies illustrate that FM can be successfully fabricated and that some patients with FM can be misidentified as malingerers. The result can be a costly error to either the unidentified or mislabeled FM sufferer, or to the insurer or tort defendant if a claimant successfully fabricates or exaggerates disability associated with FM.

**Prevalence & Gender**

Several studies have investigated the prevalence rates of FM in the general as well as patient populations. For example, Croft, Rigby, Boswell, Schollum, and Silman (1993) found the prevalence of chronic widespread pain, the type of pain reported in FM patients, was 11.2% in a sample of the general population in the North of England. An epidemiological study in Canada estimated that 5% of women and 1.6% of men in one eastern city have FM (White, Speechley, Harth, & Østbye, 1999). Other studies have estimated that three to six million people (i.e., approximately 1-2%) suffer from FM in the United States (Campbell et al., 1983; Goldenberg, 1987). Studies conducted in Canada and the United States found that approximately 10% of the general population has widespread pain (pain characteristic of FM; White et al., 1999; Wolfe, Ross, Anderson, Russell, & Hebert, 1995). A more detailed examination of those with widespread pain indicates that approximately 2% of the general population would meet the diagnostic criteria for FM set out by the ACR (Wolfe et al., 1990). In medical settings, FM is one of the most common disorders seen in outpatient rheumatology clinics (Goldenberg, 1987; White, Speechley, Harth, & Østbye, 1995; Wolfe et al., 1995).

FM is 5 to 10 times more frequent in women than men (Bennett, Clark, Campbell, & Burckhardt, 1992). Wolfe et al. (1995) found that 80 to 90% of those afflicted with FM are women, and that their mean age is in their mid-forties. These authors reported that prior to adulthood, the gender distribution is approximately equal for boys and girls. One study revealed that adult males were less likely to have lower body tenderness than upper body tenderness, but their likelihood of meeting the FM criteria was significantly increased when lower body tenderness was present (White et al., 2000). In a review of gender and FM symptoms, Yunus (2001) reported that women tended to experience significantly more
severe and a greater number of FM symptoms, including widespread pain, more tender points, fatigue, and more problems with physical functioning, than men.

A fibromyalgia-like syndrome has existed for many years, under various names such as fibrositis, myofascial pain, muscular rheumatism, psychogenic rheumatism, and tension myalgia (Waylonis, Ronan, & Gordon, 1994). Historically, some of those names (e.g., psychogenic rheumatism and tension myalgia) have been associated with medical views that imply that FM is psychiatric in origin. Once a particular perception has been established in the medical community, it can be difficult to alter even with compelling contrary evidence. To illustrate, Monmaney (1993) described the difficulties encountered in persuading physicians to consider evidence that peptic ulcers were an infectious disease caused by bacteria, and not a psychosomatic disease. Physicians were of the belief that peptic ulcers were the result of stress or unmet needs, their beliefs influenced the diagnosis rather than an unbiased consideration of alternate information. Lidbeck (2002) suggested that in the absence of objective findings, patients’ descriptions of pain and observed pain reactions are often interpreted as ‘nonorganic signs’, or as features of ‘psychogenic’ pain.

It may be theoretically possible that a gender bias has influenced both the rate of diagnosis and the controversy regarding whether the etiology of FM is physical, psychological, or both. That is, women are much more likely to be diagnosed with the condition, and women have, historically, been more susceptible to psychogenic explanations for physical conditions. Vertinsky (1990) provided the opinion that the treatment of women by physicians in the late 1800s was a form of social control. The common diagnoses of hysteria and neurasthenia were given to women who showed signs of fatigue and anxiety. Those historical diagnosing patterns suggest that the medical profession considered mental instability rather than a physical illness more frequently for their female patients. Those diagnosing patterns further suggest an increased possibility that FM will be perceived as psychogenic in origin, if women report it more frequently.

It has been reported that some women have been misdiagnosed with psychosomatic illnesses, even in the modern era when an illness or disease has a well-defined biological basis. For example, Russell (1985) found that in a sample of 21 women and 14 men with multiple sclerosis in Canada, seven women and one man were told by their physicians that their symptoms were psychological in origin. A more recent study examining patterns of
Credibility and Fibromyalgia

misdiagnosis similarly reported that a disproportionate number of women with multiple sclerosis were told by their physicians that they suffered from psychiatric problems or medically unexplained problems (Levin, Mor, & Ben-Hur, 2003). Issues and concerns regarding misdiagnosis in women with autoimmune diseases, such as systemic lupus erythematosus, continue to be expressed (Reeves, 2004). Several authors have provided anecdotal evidence that women are more likely to receive a psychiatric diagnosis on the basis of physical symptoms (Hannaford, 1985; Jefferys, 1982; Ramsay, 1986; Register, 1987; Thorne, 1993).

The epidemiological research and diagnostic labels that have previously been used to refer to FM suggests a possible gender bias in the rate of diagnosis and controversy over the etiology of FM. Historical accounts of diagnostic labels and medical treatments related to women further indicate that women have been more susceptible to psychogenic explanations for physical conditions. These possible biases may influence the credibility of women reporting symptoms of illnesses with uncertain etiologies, such as FM. Research linking gender differences, pain sensitivity, and pain inhibitory mechanisms described in the medical section below may provide an explanation for the gender disparity of FM.

Biomedical Factors

There have been numerous attempts to identify underlying physiological problems associated with specific FM symptoms or with the syndrome in general. The research has been mixed. Some physiological correlates have been found, and replicated, but underlying pathophysiologies for specific symptoms and the syndrome in general remain elusive. Literature representing the biomedical research will be reviewed in this section.

Because the central feature of FM is widespread musculoskeletal pain, researchers have attempted to identify abnormal muscle structures or tissue that might be the cause of the pain. Research examining peripheral muscle tissue and structures of FM patients found that, although some fiber abnormality was identified (Bengtsson, Henriksson, & Larsson, 1986), it was not specific to FM and no definitive pathology was found (Drewes & Andreassen, 1993). Another study suggested that any abnormal findings were due to chronic deconditioning (Simms, 1996).

Researchers have studied the relationship between sleep disturbance and FM. Many patients with FM report sleep disturbance, wherein they fall asleep without difficulty but
Credibility and Fibromyalgia

consistently awake unrefreshed (Wolfe et al., 1990). Early research suggested that stage 4 sleep might be disrupted in some patients with FM (Moldofsky & Scarisbrick, 1976; Moldofsky, Scarisbrick, England, & Smythe, 1975). However, results from subsequent studies have not replicated these findings (e.g., Shaver et al., 1997). Stage 4 sleep is essential in the secretion of the growth hormone, insulin-like growth factor (IGF-1, formerly called somatomedin C), that plays an important role in muscle homeostasis and repair. Two studies have identified a growth hormone deficiency state in approximately one-third of patients with FM (Bennett et al., 1992; Bennett, Cook, Clark, Burckhardt, & Campbell, 1997). However, one study did not find this deficiency state in patients with FM (Buchwald, Umali & Stene, 1996).

A growth hormone deficiency in adults has been associated with symptoms similar to patients with FM, and replacement therapy has been reported to improve those symptoms, including energy level, dysthymia, and quality of life (Cuneo, Salomon, McGauley, & Sonksen, 1992; Cuneo, Salomon, Wiles, Hesp, & Sonksen, 1991a; McGauley, Cuneo, Salomon, & Sonksen, 1990), as well as exercise capacity (Cuneo et al., 1991b). Bennett, Clark, and Walczyk (1998) explored replacement therapy in women with FM, involving daily injections of a hormone identical to human pituitary derived growth hormone. These women displayed a significant decrease in the tender-point count, improved exercise capacity, and improved quality of life. However, Bennett and colleagues cautioned that the cost of this therapy prohibits use with the subset of FM patients who demonstrate this deficiency.

There has been considerable research relating to neuroendocrine functioning and FM. This is partly due to the belief that stress can initiate or exacerbate FM (Wolfe et al., 1990). Researchers have identified deficiencies in the hypothalamic-pituitary-adrenal (HPA) axis in patients with FM compared to healthy individuals (Clauw & Chrousos, 1997; Crofford et al., 1994; Griep, Boersma, & de Kloet, 1993; van Denderen, Boersma, Zeinstra, Holander, & van Neerbos, 1992). These deficiencies have been characterized by an increased adrenocorticotropic hormone (ACTH) response, and a blunted adrenal secretion of cortisol. Dessein, Shipton, Stanwix, and Joffe (2000) suggest that the neuroendocrine studies indicate deficiencies in some patients with FM that may mediate the cause and persistence of pain and associated symptoms.
A dysregulation of the HPA axis has also been identified in patients with major depression (for a review see Plotsky, Owens, & Nemeroff, 1995), and high prevalence rates of depression have been reported in patients with FM, as will be discussed in the psychology section. Specifically, some studies have reported greater ACTH levels (Kalin, Weiler, & Shelton, 1982; Reus, Joseph, & Dallman, 1982; Young, Carlson, & Brown, 2001), while other studies have reported normal to low ACTH levels in patients with depression (Fang, Tricou, Robertson, & Meltzer, 1981; Yerevenien & Woolf, 1983; Sherman, Phohl, & Winokur, 1985; Poesner et al., 2000). Parker, Schatzberg, and Lyons (2003) have suggested that the inconsistencies of ACTH levels in depressed patients are related to time, so that normal to high levels occur in acute depressions, while low levels of ACTH occur in chronic depressions. According to their explanation, the HPA axis in patients with FM more closely resembles that of the acutely depressed patient. Given the overlap of FM and depression, it is difficult to determine whether the HPA axis deficiencies in some patients with FM are associated with the FM, the depression, are a feature of their comorbidity, or are related to something else.

The roles of neurochemicals and neurotransmitters in the process of transmitting pain signals in the central nervous system (CNS) have been studied in FM. Two such biochemicals involved in pain processing are serotonin and substance P. A series of studies have consistently found lower levels of serotonin in the cerebrospinal fluid (Hrycaj, Stratz, & Muller, 1993; Russell et al., 1992; Russell, Værøy, Javors, & Nyberg, 1992; Wolfe, Russell, Vipraio, Ross, & Anderson, 1997; Yunus, Daily, Aldag, Masi, & Jobe, 1992), and in the peripheral platelets of patients with FM (Russell & Vipraio, 1994). Ernberg, Lundberg, and Kopp (2000) examined the effects of increased serotonin levels through intramuscular injection and found an increase in pain for healthy individuals, but not for patients with FM. A related investigation demonstrated that an infusion of a dorsal horn receptor antagonist, ketamine, reduced pain in the experimental treatment of patients with FM (Graven-Nielsen et al., 2000). Of course, decreased levels of serotonin have also been identified in patients with major depression (for reviews see Celada, Puig, Amargos-Bosch, Adell, & Artigas, 2004; Fava, 2003; Sobczak, Honig, van Duinen, & Reidel, 2002), and as mentioned earlier high prevalence rates of depression have been reported in patients with FM. In fact, the association between depression, pain, and the role of serotonin has been recently reviewed
Credibility and Fibromyalgia (Delgado, 2004; Stahl & Briley, 2004), emphasizing the role of serotonin in the physical symptoms of depression and pain. It is therefore difficult to determine whether low serotonin in some patients with FM is associated with the FM, the depression, or both factors.

Significantly elevated levels of substance P have been found in the cerebrospinal fluid of FM patients, levels that were two to threefold higher than in healthy controls (Liu, Welin, Bragee, & Nyberg, 2000; Russell et al., 1994; Wellin, Bragee, Nyberg, & Kristiansson, 1995; Væroy, Helle, Førre, Káss, & Terensuis, 1988), and found to be stable or increase slightly over time (Russell, Fletcher, Vipraio, Lopez, & Orr, 1998). The results of several studies indicate that excitatory amino acids are also involved in pain transmission at the spinal cord level (Cotman, Monaghan, Ottersen, & Mathison, 1987; King & Lopez-Garcia, 1993; Raigorodsky & Urca, 1990; Skilling, Smullin, & Larsson, 1988). Larson, Giovengo, Russell, and Michalek (2000) found that concentration levels of amino acids correlated with the tender point index and individual measures of pain intensity in patients with FM.

Some patients with FM perceive stress as an aggravating factor of their condition (Okifuji & Turk, 2002). The possibility of a disturbed stress-response system has been considered in patients with FM. For example, Værøy et al. (1988) found Raynaud’s syndrome (a circulatory dysfunction) was reported in a disproportionately high number of patients with FM, compared to healthy controls. Subsequent research found a reduced vasoconstrictory response to cold and auditory stress stimuli in patients with FM, compared to healthy controls (Værøy, Qiao, Morkrid, & Førre, 1989). Patients with FM have also demonstrated a lower level of blood flow during a stress test (cold pressure test) relative to healthy individuals, supporting the view of a blunted autonomic response to stressors (Qiao, Værøy, & Morkrid, 1991). FM patients had significantly lower plasma cortisol levels following exercise, compared to healthy controls, even though pre-exercise levels were similar (van Dendreren et al., 1992). These findings suggest that some patients with FM exhibit an impaired stress response to physiological stimuli, and that a dysfunctional autonomic stress response system might be associated with FM.

Psychophysical studies have been conducted to examine objective measures of pain stimuli and the threshold at which those stimuli are considered painful in patients with FM. One such study examined pain perception by applying varying levels of force with an electrical palpometer, and recorded self-reports of pain intensity on a visual analogue scale
Credibility and Fibromyalgia

(Bendtsen, Norregaard, Jensen, & Olesen, 1997). The pain threshold was twice the number of units of applied force on the palpometer for pain-free control subjects, compared to patients with FM. Other studies have demonstrated that FM patients and control subjects generally detect electrical, thermal, and mechanical sensory stimulation at the same levels, but the level at which these stimuli become noxious or unpleasant (pain threshold) is significantly lower in FM patients (Arroyo & Cohen, 1993; Gibson, Littlejohn, Gorman, Helme, & Granges, 1994; Kosek, Ekholm, & Hansson, 1996; Lautenbacher, Rollman, & McCain, 1994).

McDermid, Rollman, and McCain (1996) have explored pain mechanisms in FM through the hypervigilance model of pain perception. The model states that chronic pain patients have a heightened sensitivity to pain because of increased awareness and preoccupation with noxious stimuli. Mechanical and auditory sensory stimuli were used, and hypervigilance was measured through questionnaires designed to assess an individual’s reactivity level, sensitivity to noise, and the frequency of occurrence of physical symptoms. They found that FM patients had significantly lower pain threshold and tolerance values than rheumatoid arthritis patients, who in turn had lower values than healthy controls. The authors concluded that patients with FM and various other pain conditions may have an exaggerated focus on physical experiences. They further suggest the importance of examining the possible psychological and biological factors that may be involved in a hypervigilant pain perception, and to examine whether this response pattern is typical in patients with disorders of uncertain etiology. However, results from a subsequent study did not replicate these findings (Peters, Vlaeyen, & van Drunen, 2000). A more recent study of hypervigilance found that although patients with FM reported more vigilance to pain than chronic low back pain patients, the heightened vigilance was mediated by pain intensity and catastrophic thinking about pain (Crombez, Eccleston, Van den Broeck, Goubert, & Van Houdenhove, 2004). They suggested that FM patients may consider the origin of their pain as more mysterious than chronic low back pain, giving them ample opportunity to worry and catastrophize about their pain. This view was supported by a study that found women with FM were more prone to catastrophizing their pain than women with rheumatoid arthritis (Hassett, Cone, Patella, & Sigal, 2000). Crombez and colleagues concluded that the vigilance is not a unique characteristic of FM, but is related to pain intensity and catastrophic thinking about pain. Finally, the authors suggested that pain intensity and catastrophic thinking about
pain are likely to cause a heightened vigilance to pain, but that the opposite might also be true.

Functional neuroimaging has been used to identify brain structures involved in human pain processing. Coghill, McHaffie, and Yen (2003) used functional magnetic resonance imaging (fMRI) to identify areas of brain activation that might be associated with subjective reports of pain. Healthy subjects underwent functional imaging during the delivery of painful thermal stimuli. The cerebral cortical regions important in sensation, attention, and affect (anterior cingulate cortex, primary somatosensory cortex, and ipsilateral prefrontal cortex) showed significantly greater activation in subjects who reported higher levels of pain than those who reported lower levels of pain. In contrast, no significant differences were found in the thalamic regions, which are involved in transmitting pain signals from the spinal cord to higher brain regions. These findings suggest an association between regional activation and pain perception, and that cerebral cortical regions may play a central role in aspects of individual differences in pain perception.

Gracely and colleagues also investigated the pattern of cerebral activation with fMRI during the application of painful stimuli in patients with FM (Gracely, Petzke, Wolf, & Clauw, 2002). The subjective pain reports and brain activation of patients with FM were qualitatively and quantitatively similar to control subjects when at least twice the pressure was applied to the control subjects. These results suggest an augmentation of pain sensitivity in patients with FM at the cortical and subcortical level. Similar findings have been observed in patients with allodynia (pain experienced by non-painful stimuli) caused by cerebral infarction (Peyron et al., 1998).

The lower pain threshold that the psychophysical studies found in patients with FM led to further exploration of this phenomenon at a neural level. Mendell and Wall (1965) initially discovered that repetitious peripheral stimulation of noxious input resulted in a progressive build-up of the magnitude of the electrical response recorded in the dorsal horn neurons. This build-up or temporal summation of electrical response is referred to as "wind-up" or "second pain", and results in alldynia/hyperalgesia and spontaneous pain. This wind-up is dependent on an amino acid receptor (N-methyl-D-aspartic acid) and substance P synaptic mechanisms within the dorsal horn of the spinal cord (Dickenson & Sullivan, 1987; Price, Mao, Frenk, & Mayer, 1994; Woolf & Thompson, 1991). A pain stimulus can activate
A fibers (short-latency impulse discharges) and C fibers (long-latency discharges). Both types of fibers discharge impulses into dorsal horn neurons specific to pain reception. The dorsal horn neurons associated with pain-related pathways undergo central sensitization during activation from C fibers, critical for inducing and maintaining central hyperalgesic states that accompany persistent pain conditions, and wind-up pain.

Researchers have reported that enhanced responsiveness to painful and non-painful stimuli occur after sustained input from a painful stimulus in healthy individuals (for a review see Coderre, Katz, Vaccarino, & Melzack, 1993). Wind-up is often abnormally triggered or exaggerated in pain conditions such as complex regional pain syndrome (Price, Bennett, & Rafii, 1989) and post-herpetic neuralgia (Eide, Jorum, Stubhaug, Bremnes, & Breivik, 1994). Staud, Price, and colleagues have explored the possibility that wind-up may help explain some characteristics of FM pain. Heat taps (contact with a preheated metal probe) were delivered to the palm of either hand for 700 ms to induce wind-up pain, and subjects rated pain on a visual analogue scale. Specifically, they found differences in response to repetitive thermal stimuli, so that perceived magnitude of secondary pain in response to the first stimulus was greater, and levels of temporal summation and after-sensations were greater for FM than control subjects (Staud, Vierck, Cannon, Mauderli, & Price, 2001). They found a similar pattern of results in response mechanical stimulation of muscle tissue (Staud et al., 2003). Price et al. (2002) examined whether enhanced wind-up could be modulated by central inhibitory mechanisms in patients with FM. They found that wind-up related to C fiber transmission was attenuated by placebo and analgesic injections in both FM and control subjects, but more immediate pain sensations related to A fiber transmission was not attenuated in either subject group. They reported that the lack of blockade of placebo effects by the analgesic may have been due to an inadequate drug dose. The authors further conclude that the results may indicate that some central pain-modulatory mechanisms are not deficient in patients with FM. Additional research is necessary to determine whether a psychological component may have been involved in this outcome.

Testing the responsiveness of central pain inhibitory mechanisms has been investigated through a paradigm used to examine diffuse noxious inhibitory controls (DNIC). That paradigm involves the exploration of pain responsiveness at multiple sites following the experience of pain at one site. Results have shown that strong sustained pain decreases pain
Credibility and Fibromyalgia

responsiveness at other locations in healthy individuals (Price & McHaffie, 1988; Talbot, Duncan, Bushnell, & Boyer, 1987). Results from two studies examining inhibitory pain mechanisms suggest a deficiency in a pain-inhibitory mechanism (DNIC) for patients with FM, but not healthy controls (Lautenbacher & Rollman, 1997; Kosek & Hansson, 1997). Related research has suggested that exercise decreases wind-up pain in normal subjects, and it increases wind-up pain in patients with FM (Vierck et al., 2001). These results support the view that modulation of pain sensitivity is compromised in FM. Because FM is more prevalent in women than men, Staud, Robinson, Vierck, and Price (2003) examined whether there are gender differences in DNIC, or whether this inhibitory control was more characteristic of FM per se. A gender effect was reported, indicating that healthy women and women with FM were generally lacking this pain-inhibitory mechanism, compared to healthy men. The mechanism was not specific to FM.

A consistent sex difference finding in the pain modulation research is that women generally have an increased sensitivity (lower pain thresholds and tolerance levels) in experimental pain paradigms as compared to men (Edwards et al., 1999; Ellermeier & Westphal, 1995; Fillingim et al., 1999; Lautenbacher & Rollman, 1993; Maxiner & Humphrey, 1993; Riley, Robinson, Wise, Myers, & Fillingim, 1998; Rollman, 1997; Rollman & Harris, 1997). For example, gender differences in temporal summation have been reported for healthy females (Fillingim, Maixner, Kincaid, & Silva, 1998). Results indicated lower pain threshold, pain tolerance, and greater temporal summation of thermal pain for women, compared to men, but no gender differences in ability to detect small increments in painful stimuli or magnitude estimates of discrete noxious stimuli. These differences, however, have not been consistently replicated in clinical settings (Robinson, Wise, Riley, & Atchison, 1998), or in a survey of spinal cord injured patients (Cardenas, Bryce, Shem, Richards, & Elhefni, 2004).

Wise and colleagues examined gender role expectations on experimental pain reports, and found significant sex differences for measures of pain threshold, pain tolerance, and pain unpleasantness ratings (Wise, Price, Myers, Heft, & Robinson, 2002). Anxiety sensitivity, or the fear of anxiety-related sensation, is another psychological factor related to higher pain sensitivity in women (Keogh & Birkby, 1999). Furthermore, sex differences have also been found with regards to pain coping strategies (Keogh, Hatton, & Ellerby, 2000; Keefe,
Lefebvre, Egert, Affleck, Sullivan, & Caldwell, 2000; Osman, Barrios, Gutierrez, Kopper, Merrifield, & Grittman, 2000). One study examining the relationship between lifetime pain ratings and response to cold pressor pain found that individuals with greater lifetime pain were more reactive to the experimental pain, and that previous pain experience and self-reported pain tolerance was highly predictive of laboratory pain tolerance (Rollman, Abdel-Shaheed, Gillespie, & Jones, 2004). The results also revealed that women reported significantly greater lifetime pain levels from illness and injury, than men. These authors suggested that lifetime differences in pain experience combined with the gender differences in pain tolerance described above, could indicate that women have greater sensitivity to pain earlier in life, and consequently, may experience more pain than men. Hence, both psychosocial factors and biological sex differences should be considered as possible explanations for the contrasting reports between genders.

Melzack and colleagues suggest that the CNS can be transformed to provide for continuing pain without evidence of physical pathology, specifically referring to “neuroplasticity” research (Melzack, Coderre, Katz, & Vaccarino, 2001). Neuroplasticity, in the pain literature, is described as the capacity of neural structures to change their function, neurochemical profile (e.g., transmitters, ion channels, protein receptors), or structure in response to injury or other pathological events, which subsequently can affect pain sensitivity (Woolf & Salter, 2000). The resulting structural abnormalities are a type of remodeling of the neural system, including possible cellular and molecular changes, alterations to synaptic connections, and could involve changes to the spinal cord and brain structures associated with pain transmission. For example, stimulation of the cortex in people with paraplegia can produce sensations in anesthetic body areas below the spinal injury site, suggesting changes in the input, but not the output pattern of these cells (Cohen, Topka, Cole, & Hallett, 1991). In terms of chronic pain, the somatosensory cortical representation of the lower back appears to change and reorganize in some patients with chronic low back pain (Flor, Braun, Elbert, Birbaumer, 1997). Cortical reorganization appears to occur in some patients with tinnitus (Muhlneckel, Elbert, Taub, & Flor, 1998), and in amputees (Elbert et al., 1997; Flor et al., 1998, Grusser et al., 2001; Grusser et al., 2004; Karl, Birbaumer, Lutzenberger, Cohen, & Flor, 2001).
Credibility and Fibromyalgia

The neural structural changes may ultimately lead to a central hyperexcitability and sensitization, which may partially underlie the pain associated with FM, as well as other injury and non-injury associated chronic pain conditions (e.g., post-traumatic headache, myofascial pain, and whiplash pain syndromes). Bennett (1996) suggested that the continuous flow of noxious impulses that FM patients experience from musculoskeletal pain and other conditions associated with FM (e.g., skin tenderness, headaches, irritable bowel syndrome) could be the perpetuating factors in the maintenance of neuroplastic changes, resulting in its chronicity. The neuroplasticity research is still in the early stages of exploration, so more work is necessary before conclusions can be drawn regarding its relevance to FM.

Psychological Factors

Psychiatric disorders and psychological problems have received a lot of attention in the literature relating to FM. Some specific psychological problems have been associated with FM, but none that would provide a conclusive explanation for the syndrome. The psychological considerations and some possible theoretical models that help explain the experience of pain associated with FM will be reviewed in this section. Some of these psychological factors, in relation to the pain perception and sensitivity literature, were discussed in the previous section (Biomedical Factors).

The incidence of major depression reportedly is approximately doubled for individuals with chronic medical conditions (Ohayon & Schatzberg, 2003; Patten, 2001), including FM (Gagnon & Patten, 2002). Higher prevalence rates of anxiety and depressive disorders have been found among patients with FM compared to healthy controls and individuals with other rheumatic conditions (e.g., Goldenberg, 1986; Hassett et al., 2000; Hudson, Goldenberg, Pope, Keck, & Schlesinger, 1992; Hudson, Hudson, Pliner, Goldenberg, & Pope, 1985; Walker et al., 1997a). Patients with FM not only show higher prevalence rates of depression, anxiety, low energy, and poor sleep, but they also demonstrate severe functional difficulties (Borman & Celiker, 1999; Epstein et al., 1999). Some researchers, however, have not reported a difference in anxiety and depression among patients with FM when compared to patients with rheumatoid arthritis (Ahles, Yunus, & Masi, 1987, Kirmayer, Robbins, & Kapusta, 1988; Kurtze, Gundersen, & Svebak, 1998). A meta-analytic review found higher prevalence rates of depression, but not higher rates of
Credibility and Fibromyalgia

anxiety in patients with FM, compared to other patient groups and healthy controls (Henningsen, Zimmermann, & Sattel, 2003).

Cognitive factors play an important role in chronic pain conditions and how individuals adapt or cope with symptoms (e.g., Turk & Rudy, 1986). Cognitive factors refer to beliefs and perceptions an individual may have about their symptoms and abilities. Research examining the effects of maladaptive cognitions in chronic pain patients revealed that dysfunctional perceptions and interpretations of symptoms are associated with a reduction of activities, a decreased sense of accomplishment and social reinforcement, as well as further physical de-conditioning. Although the research is limited, perceptions of uncontrollability, belief that behaviours do not significantly impact on one’s condition, and self-efficacy beliefs have been shown to be related to reports of greater pain, disability, and depressive mood in patients with FM (Buckelew, Murray, Hewett, Johnson, & Huyser, 1995; Turk & Okifuji, 1997). In fact, Turk and Okifuji found that patients’ perceptions of their disability and control over their lives are closely associated with the number of physician visits, so that negative perceptions are associated with an increase in physician visits. These psychosocial issues are not unique to FM, of course. They have also played an important role in assessing disability in patients with rheumatoid arthritis and osteoarthritis (Dekker, Boot, van der Woude, & Bijlsma, 1992; Fifield, Reisine, & Grady, 1991). It should be noted that Turk, Okifuji, Starz, and Sinclair (1996) also found that approximately one-third of patients with FM seem to adapt well to their conditions despite the persistent pain and related symptoms.

Although not as commonly documented as other problems associated with FM, patients frequently complain about concentration and memory difficulties (e.g., Suhr, 2003). Performance on a number of attentional and memory tasks were examined and findings indicated that although individuals with FM performed more poorly than healthy controls, their task performance was within the range of “normal” limits across all tests (Grace, Nielson, Hopkins, & Berg, 1999). One study reported that FM patients performed more poorly than age-matched controls on a number of memory and cognitive processing tasks, with the exception of speed of processing (Park, Glass, Minear, & Crofford, 2001). Results of this study further showed that poor performance on cognitive tasks correlated with subjective pain complaints and not with depression or anxiety, in patients with FM. Sletvold,
Stiles, and Landro (1995) compared FM and depressed patients and found that both groups processed information less effectively than controls on mental tasks and psychomotor performance tasks. In a follow-up study, Landro, Stiles, and Sletvold (1997) found that when depressive status was accounted for in patients with FM, only the subsample with a lifetime major depressive disorder performed more poorly on memory tasks, compared with healthy controls. Dick, Eccleston, and Crombez (2002) compared FM, rheumatoid arthritis, and musculoskeletal pain patients with healthy controls on tasks measuring attentional functioning and found that each patient group displayed significant deficits (scores in the "clinically impaired" range) on tasks of selective and sustained attention, and auditory-verbal working memory, compared to healthy controls. Contrary to Park et al. (2001), no association was found between mood, pain related measures, and attentional functioning. Although patients with FM displayed a significantly higher level of anxiety than the other three groups, and greater somatic awareness than rheumatoid arthritis patients, they did not display greater attentional deficits than the other pain patients.

Suhr (2003) assessed the relationship of depression, pain, and fatigue to objective cognitive impairment in patients with FM, patients with other chronic pain disorders, and healthy controls. After controlling for depression, pain, and fatigue, there was no objective evidence of impairment on cognitive tests across groups, and no group differences in the number of participants who scored in the clinically impaired range on the neuropsychological test battery. Self-reported depression was significantly related to performance on a number of memory tasks, and self-reported fatigue was significantly related to psychomotor speed, which supports an earlier finding by Cote and Moldofsky (1997). Suhr (2003) also found that patients with FM reported significantly more depression, pain, fatigue, and cognitive complaints compared to other chronic pain patients and healthy controls. The findings showed that patients with FM had a greater perception of cognitive impairment than was evident in their neuropsychological test results, thus emphasizing the importance of separating subjective cognitive complaints from objective neuropsychological test performance.

Neuroendocrine levels have been associated with memory functioning, such that extremely low or high cortisol levels have been associated with reduced memory performance (Golier & Yehuda, 1998; Kirschbaum, Wolf, May, Wippich, & Hellhammer,
Credibility and Fibromyalgia

1996; Lupien et al., 1997), while moderate elevations enhance memory (Buchanan & Lovolla, 2001). A recent study by Sephton et al. (2003) examined the relationship between memory performance, neuroendocrine level, and psychological factors in patients with FM. There was a positive association between higher cortisol levels and performance on a variety of memory tasks, while self-reported depressive symptoms were negatively associated with long-term memory performance. The authors concluded that both cortisol level and depressive symptoms may play a role in the cognitive deficits demonstrated by patients with FM.

The influence of personality and personality disorders has not yet been empirically studied in the FM patient population, although it has been examined in other chronic pain populations. Research examining personality disorders among chronic pain patients reveal prevalence rates from 31% to 81% (Burton, Polatin, & Gatchel, 1997; Fishbain, Goldberg, Meagher, Steele, & Rosomoff, 1986; Large, 1986; Polatin, Kinney, Gatchel, Lilo, & Mayer, 1993; Reich, Tupin, & Abramowitz, 1983). The most frequent personality disorders found among these samples include histrionic, dependent, paranoid, and borderline. Patient sample (e.g., back pain, carpal tunnel) and type of diagnostic interview (semi-structured vs. structured) varied across these studies, which may contribute to the lack of consistency in overall rates of personality disorders. Two more recent studies have incorporated a structured diagnostic interview to examine the prevalence rate of personality disorders in chronic pain populations. The first study examined a sample of chronic low back pain patients, and results revealed a broad range of personality disorders in the pain patients relative to the normal comparison group (Vittengl, Clark, Owen-Salters, & Gatchel, 1999). The second study consisted of patients with work-related chronic musculoskeletal pain disability and results revealed a 70% prevalence rate of personality disorders in the study sample (Dersh, Gatchel, Polatin, & Mayer, 2002). The most prevalent personality disorders in this sample were Paranoid, Borderline, and Histrionic.

An association between physical and sexual abuse during childhood and/or adulthood and chronic pain syndromes has been reported (Bailey, Freedenfeld, Kiser, & Gatchel, 2003; Green, Flowe-Valencia, Rosenblum, & Tait, 2001; Goldberg & Goldstein, 2000; Lampe et al., 2003; Wurtele, Kaplan, & Keairness, 1990). One study found that although there was an association between the incidence of childhood abuse and patients with chronic pain, the type
of chronic pain (i.e., low back, craniofacial, widespread) was not significantly related to
childhood abuse (Goldberg, 1994). Another study indicated that women with a history of
childhood sexual abuse reported a greater number of painful body areas, more diffuse pain,
and were more likely to receive a diagnosis of FM (Finestone et al., 2000). Walker and
colleagues reported that patients with FM had significantly higher prevalence rates of sexual,
physical, and emotional abuse than rheumatoid arthritis patients (Walker et al., 1997b).
Imbierowicz and Egle (2003) also found that patients with FM had significantly higher
prevalence rates of childhood adversities (including childhood sexual and physical abuse)
that patients with medically explained chronic pain, and that the experiences of FM patients
were similar to patients with somatoform pain disorders. Two studies observed that women
with FM and a history of sexual abuse reported significantly more symptoms than women
with other rheumatic diseases or healthy controls (Boisset-Pioro, Esdaile, Fitzcharles, 1985;
Taylor, Trotter, & Csuka, 1995). However, these investigators did not find a higher
prevalence rate of childhood sexual abuse between women with FM and the other groups.
Van Houdenhove et al. (2001) only found higher prevalence rates of emotional neglect and
physical abuse, in patients with FM, but not sexual abuse. Although there appears to be an
association between childhood abuse history and both chronic pain and FM, the nature of this
relationship is unclear.

The biomedical model has been widely used to understand disability. One of the main
principles of this model is that disability is primarily attributable to an underlying organic
pathology (Engel, 1977). Another central principle of this model is the fundamental
separation between the physical and mental, and a de-emphasis on the role of psychosocial
factors in disability (Schultz, Crook, Fraser, & Joy 2000; Turk & Okifuji, 2002). It has been
suggested that this model applies well to individuals with an acute or uncomplicated injury,
or individuals suffering from serious medical conditions (Schultz et al., 2000).

The insurance model, also known as the forensic model, views disability claimants
who anticipate secondary gains as likely candidates to be dishonest about their condition
(Frank, Pulcins, Kerr, Shannon, & Stansfeld, 1995). According to this model, the primary
focus is on differentiating between “honest” and “dishonest” claimants. The insurance model
shares the biomedical model’s need for objective evidence of an organic pathology.
The biopsychosocial model has been proposed in the understanding of disability. It was first formulated by Engel (1977), who proposed that illness represented a complex interaction of biological, psychological, and social influences, with each component affecting and being affected by the other. For example, back pain may initiate a psychological/emotional reaction of frustration or anger. This reaction may be affected by the person's social environment; that is, a spouse or friends may be overly solicitous and provide excessive attention when the person reports their pain. The social responses may lead the person to avoid movement and exercise. This can result in muscle stiffness and weakness and, hence, perpetuate the pain. The biopsychosocial model applies to chronic pain conditions, and the interaction of factors that influence their functioning.

The labour relations model has been applied to the diagnosis and rehabilitation of pain-related occupational disability, and shares the multidisciplinary approach to disability that is central to the biopsychosocial model. One of the main principles of this model is that disability is understood in an employment context, and that personal, economic, physical, psychological, social, and domestic factors are involved in the disability process (Shrey & Olsheski, 1992).

Banks and Kerns (1996) have proposed a diathesis-stress model to explain the high rates of depression commonly associated with chronic pain patients. The model asserts that underlying genetic and environmental predispositions become expressed in stressful situations. Weisberg and Keefe (1997) have applied this model to personality disorders in the chronic pain population, and suggested that personality patterns associated with marginally adaptive coping styles prior to the onset of pain, typically decompensate with the stress of an injury, chronic pain, or disability, resulting in a personality disorder. Dersh, Polatin, and Gatchel (2002) support this view and added that, while the stress associated with pain exacerbates the predisposition of psychopathology in some individuals, that psychopathology intensifies the pain experience in a reciprocal manner. Examining personality disorders in patients with FM could provide us with a better understanding as to why some individuals with FM function well, while others struggle (Turk et al., 1996).

Because the etiology and pathology of FM have not yet been identified, there is no way of determining whether FM is a separate entity or clusters of symptoms that do not share an originating source. There are examples of inconsistent findings related to FM in the
medical and psychological research to support this view. Okifuji and Turk (2002) provide an explanation of the inconsistent findings by proposing that there may be subsets or subgroups of patients within the general diagnostic criteria of FM. Their physiologic-heterogeneity hypothesis is based on previous research examining psychosocial and behavioural characteristics of FM patients (Turk et al., 1996). They further postulate the possible existence of subgroups based on immunologic, endocrinological, and autonomic nervous system disturbances. The brain imaging research supports the individual differences hypothesis with the view that cerebral cortical regions may play a central role in aspects of individual differences in pain perception (Coghill et al., 2003). Clearly, research designed to look at individual differences in the experience of pain, and explore the possible existence of subgroups within FM, could be useful from both medical and psychological perspectives to further understand this issue.
Summary

Prevalence & Gender

The overwhelming disparity of female over male FM patients raises the question of why a gender disparity exists. A partial explanation may be related to the findings that women generally were found to have an increased pain sensitivity (lower pain thresholds and tolerance levels) in experimental pain paradigms as compared to men (Edwards et al., 1999; Ellermeier & Westphal, 1995; Fillingim et al., 1999; Lautenbacher & Rollman, 1993; Maxiner & Humphrey, 1993; Riley et al., 1998; Rollman, 1997; Rollman & Harris, 1987). Pain modulation research also suggests a gender effect, with healthy women and women with FM generally lacking a pain-inhibitory control mechanism (Staud et al., 2003). The tendency for women to report a greater pain history has also been suggested to play a role in the gender differences related to pain tolerance and sensitivity (Rollman et al., 2004). Gender role expectations differed on pain reports for measures of pain threshold, tolerance, and unpleasantness ratings (Wise et al., 2002). Furthermore, sex differences were also reported for pain coping strategies (Keogh et al., 2000; Keefe et al., 2000; Osman et al., 2000) as well as anxiety sensitivity to pain stimulus (Keogh & Birkby, 1999). Rollman et al. (2004) suggests that both biological and psychosocial factors contribute to the explanation of pain related gender differences.

Of course, the clinical literature indicates that a disproportionate number of women suffer from several acute, recurrent, and chronic pain syndromes, including chronic fatigue syndrome, musculoskeletal pain, oral-facial pain, and post-surgical pain (Dao & Le Resche, 2000; Heitkemper & Jarrett, 2001; Jason et al., 1999; Morin, Lund, Villarroel, Clokie, & Feine, 2000; Robinson et al., 1998, Rollman & Lautenbacher, 2001). A predominantly female sex difference also extends to numerous medical diagnoses, including several thyroid (Hashimoto’s thyroiditis, Graves’ hyperthyroidism), rheumatic (systemic lupus erythematosus, Sjögren’s syndrome, rheumatoid arthritis), and hepatic diseases (autoimmune hepatitis, primary cirrhosis; for a review, see Lockshin, 2001). Biological and environmental differences are reported as possible factors responsible for the predominantly female sex difference in the medical diagnoses mentioned above (Lockshin, 2002). Psychosocial, biological, environmental, or some combination of factors might also account for the higher
prevalence of FM in women. Further research is needed to examine whether women present any predisposing factors in the development of FM.

**Medical Research**

There is currently no evidence supporting specific histologic changes within muscles other than what might be explained through deconditioning. However, researchers investigating sensory processing have reported significantly lower pain thresholds, greater pain sensitivity, and slower speed of recovery from noxious stimulation in patients with FM, compared to healthy controls.

The psychophysiological research suggests an impaired stress response and dysfunctional autonomic stress response system in patients with FM. However, the methodology related to this research was limited in that it did not examine how FM patients respond to personal stressors, as opposed to the standard stress tests used in experimental settings. In addition, it is difficult to determine whether the autonomic stress response system in patients with FM are a result of FM alone, or are a feature of the reported comorbidity with anxiety and depression.

The brain imaging research not only supports the view of a lower pain threshold in FM, but also suggests that cerebral cortical regions may play a central role in aspects of individual differences in pain sensitivity. Some of the sensory processing findings suggest gender differences in pain sensitivity as well. Furthermore, there appears to be growing support that the pain associated with FM is related to neurohormonal differences, neurochemical changes in the CNS, or alterations in regional blood flow in the brain. Some of the abnormalities found in FM, namely low levels of serotonin and elevated levels of ACTH and substance P are consistent with the decreased pain threshold and enhanced responsiveness patients with FM demonstrate to painful stimuli. However, low levels of serotonin and dysregulation of ACTH levels have also been reported in patients with major depression. Given the high prevalence rates of depression reported in patients with FM, it is difficult to determine whether the neurohormonal and neurochemical differences in patients with FM are simply a feature of the comorbidity.

Upon reviewing the medical findings, Bradley (1998) and Staud (2004) have concluded that there is a growing body of evidence to suggest that FM is a disorder characterized by abnormal CNS processing of sensory or nociceptive (pain) stimuli. Späth
Credibility and Fibromyalgia

(2002) shares this view and regards FM as a pain amplification syndrome. Similarly, Russell (2002) expresses the view that FM represents a human model of allodynia. Crofford and Clauw (2002) suggest that in addition to neurobiological mechanisms, behavioral factors also play a role in the expression of symptoms in many patients with FM. Alternatively, McDermid et al. (1996) suggest a hypervigilance model of pain perception to account for FM. Rollman and Lautenbacher (1993) suggest that FM is more consistent with psychophysiological formulations that involve both physical and psychological factors. As well, Okifuji and Turk (2002) have considered the possibility of a disturbed stress-response system for patients with FM. There remains no clear evidence of etiology or precipitating events for FM. The medical research findings are in need of further clarification and are far from conclusive regarding the pathogenesis of FM.

Psychology Research

Patients with FM have been shown to have significant and diverse psychological and psychiatric problems, including mood and anxiety disorders. Lepine and Briley (2004) examined the association between pain and depression, and conclude that there is no specific cause-effect relationship, because there is evidence to suggest that pain can be a risk factor for the onset of depression, and vice versa. Dysfunctional belief systems and perceptions have been associated with decreased physical functioning and increased sense of disability. The cognitive processing research indicates reduced performance on memory and attentional tasks for patients with FM, with some studies also reporting a relationship between cognitive deficits and self-reported pain, depression, and fatigue. However, results further showed no evidence of poor cognitive performance after controlling for mood and anxiety. The neuroendocrine research suggests a relationship between cortisol levels, depression, and cognitive performance in patients with FM. The use of medications must be considered in the cognitive research, because it can complicate the interpretation of results in cognitive test performance. For example, older antidepressant medications have been associated with reduced memory performance, and to a lesser extent with some of the recent medications (for a review see Cassano & Fava, 2004). Although personality disorders have not yet been examined in an FM population, results with other chronic pain populations suggest that maladaptive personality patterns and coping styles may be at work for some patients with FM.
Although there appears to be an association between childhood abuse history, and both chronic pain and FM, the nature of this relationship is unclear. Some of the studies are well designed but others have significant methodological limitations. For example, the Finestone et al. (2000) reported findings “representative” of patients with FM, but only 6 of the 80 study participants were considered to have FM, and the diagnosis of FM was provided on a self-report basis alone. As well, one study did not incorporate a control group (Green et al., 2003).

The American Medical Association (AMA) provides guidelines for the evaluation of permanent impairment (2001). The guidelines define impairment in terms of the loss of an organ, system, or function, whereas disability is defined in terms of decreased functional capacity of a personal, social, or occupational nature. The World Health Organization (World Health Organization, 1997) defines impairment in terms of the loss of physical functioning of a body system or part, while disability is defined in terms of physical, personal, and social functioning. Schultz et al. (2000) indicate that the biomedical and insurance models ignore the multidimensional nature and variety of reactions to pain. The labor relations model tends to apply more to workplace disability situations, while the diathesis-stress model focuses mainly on coping styles and personality patterns. Schultz et al. (2000) indicate that the biopsychosocial model is more applicable to chronic conditions, and in cases where complicating psychosocial and other clinical factors are involved. Both the biopsychosocial and physiologic-heterogeneity models suggest that a variety of factors should be considered in understanding how a person experiences and expresses pain, as defined by the AMA’s and World Health Organization’s definition of disability. Biological, psychological, personality, and social influences could be involved in the experience of pain, as well as the additional stress that may be involved in chronic pain conditions. These factors are important in understanding FM and the possible influences that can effect coping styles in patients with FM, especially when patients find themselves in stressful situations dealing with the disability claims and compensation systems.

Various problems associated with FM have been summarized in the sections above, and indicate that FM can be very disruptive to the patients’ quality of life, not to mention the associated increase in healthcare costs and disability claims. These issues tend to be very complex and costly from the insurance claims perspective, as will be presented below.
Credibility and Fibromyalgia

Study 1: FM in the Canadian Courts

The focus of the first study involves not only an examination of the issues faced by patients with FM and the medical system, but also issues faced by the legal system and insurance industry. To appreciate the factors and challenges associated with FM claims from the legal and insurance industry perspectives, a basic understanding of the components that comprise these systems will be presented in the following section. First, a section reviewing issues related to the private and public insurance industry and disability claims will be presented, followed by perspectives of the insured and the insurer and the inherent conflict of interest involved in those perspectives. The legal context of disability law will also be presented, followed by factors related to proof of disability and expert evidence. Issues of legal damages related to FM and disability claims will also be reviewed, along with an explanation of fraud and issues of misrepresentation.

Insurance and Disability Claims Issues

Both private and public disability insurers view pain as an immense problem (Beaty, 2002; DiBlase, 1988; Goldhammer & Bloom, 1983). Subjective symptoms such as pain have created more controversy within the disability determination process than any other issue (Rodgers, 1991). Individuals with FM reportedly have reduced work capacity and general functional disability (Hawley & Wolfe, 1991; Wolfe, 1993). They may have disability issues that are comparable to persons with rheumatoid arthritis (Cathey, Wolfe, & Kleinheksel, 1988). Buchwald and Garrity (1994) found that the mean number of health care visits per year was 22 for patients with chronic fatigue syndrome and 39 for patients with FM; 30% and 57%, respectively, were disabled from working by their illness. Similarly, a study comparing chronic fatigue and FM found an average of 21 health care visits per year for each patient group, and rates of unemployment for one year ranged from 37% for chronic fatigue, 36% for FM, and 51% for patients with both chronic fatigue and FM (Bombardier & Buchwald, 1996). It has been estimated that approximately 30% of FM patients accept shorter work hours or less physically demanding work to maintain employment, and approximately 15% currently receive disability funding associated with their symptoms (Wolfe et al., 1997b). In a related study, Wolfe et al. (1997c) calculated annual direct medical costs incurred by each FM patient to be an average of $2,274 (US dollars).
The number of claims being made against disability insurance companies has been increasing over time. Cavanaugh (1999) noted that the total dollar value of individual disability benefits paid in Canada was 240,000,000 compared with 168,000,000 in 1992. From a review of long-term disability (LTD) claims at the Canadian insurance company London Life, Cameron (1995) reported that FM was found in 11.2% of musculoskeletal claims, representing 2.8% of all group LTD claims. In addition, estimates indicated that the Canadian LTD carriers would pay out $46 million in one year for FM claims. Several longitudinal studies suggest that despite various forms of treatment, symptoms and problems remain stable over the years (Bengtsson, Bäckman, Lindblom, & Skogh, 1994; Felson & Goldenberg, 1986; Hawley, Wolfe, & Cathey, 1988; Ledingham, Doherty & Doherty, 1993; Mengshoel & Haugen, 2001; Wigers, 1996). Results from these studies suggest that FM is associated with considerable personal and occupational disability, high rates of unemployment, and a large economic burden.

The controversy continues in the insurance industry as companies struggle to determine whether to cover disability claims for conditions such as FM. Gemignani (1999) summarizes one FM claimant’s fight for benefits with the involved insurers, and describes a number of insurers and their affiliated medical advisors (physicians) who do not believe that FM is a valid, disabling medical condition worthy of insurance benefits. The denial of benefits is fueled by the fact that FM criteria are based on self-reports and there is no definitive way to determine whether someone is disabled from FM. As a result, some insurance companies refuse to pay for FM (Schlechter, 1999). The issue of credibility becomes central in these disability determinations.

The Insurance Industry

The perspectives of the insured and the insurer can be extreme and involve an inherent conflict of interest. The conflict intensifies when disability determinations involve claims for undetermined illnesses such as FM, and the issue of credibility becomes central. The perspectives of the insured and the insurer will be described in this section.

According to Ericson, Barry, and Doyle (2000), Canadian insurance industry associations frequently proclaim that claims fraud is the second leading source of criminal profits next to drugs, and that 10 to 20% of all claims are fraudulent. These authors suggest that the insurance industry figures have been exaggerated in order to justify the considerable
expansion of investigative units in the past few years. Ericson and colleagues report that while the industry claims it is simply responding to an increasing problem of fraud, one industry executive indicated that it has resulted from industry pressure to "keep premiums down." The companies incorporating Special Investigation Units use surveillance systems to examine differences among individuals to (a) form more precise risk assessments, (b) develop policies aimed at particular populations, and (c) ultimately increase profit for the insurance company (Simon, 1987). Ericson described insurance industry efforts in progress to detect fraudulent claims by accessing assistance from co-workers, neighbours, and others through informant-tip lines. Healthcare professionals are also being directed to keep claims costs down, and are provided with instructional materials on how to detect fraudulent claims and malingering patients.

Of course, the industry has good reason to be concerned about rising costs. For example, Canada spent more on healthcare than on education by 1998, and there is an increase scheduled for future spending (Simpson, 2000). The cost of healthcare in the United States shows a 200-fold increase over the past thirty years (O'Brien, 1999). Rising healthcare costs fuel a rise in insurance premiums, and the insurer must attempt to prevent their costs from increasing.

Fraud is a source of increasing costs to the insurer, and researchers have demonstrated that certain conditions, such as FM, can successfully be feigned (e.g., Khostanteen et al., 2000). Similarly, Rogers (2002) has indicated that 15% of forensic cases and 7% of non-forensic evaluations are feigning a mental disorder, and that the majority of malingers are not detected by clinicians. In a study involving 333 individuals claiming compensation for hearing loss, the incidence of exaggeration on hearing tests was 17.7% (Rickards & De Vidi, 1995). In a review of studies examining malingering and dissimulation, malingering was reported to be present in 1.25 – 10.4% of chronic pain patients (Fishbain, Cutler, Rosomoff, & Rosomoff, 1999). However, the authors cautioned that because of either methodological flaws or possible referral source biases, the prevalence rates are not reliable. A more recent study reported base rates for malingering and symptom exaggeration of 35% for a combination of FM and chronic fatigue syndrome patients (Mittenberg, Patton, Canyock, & Condit, 2002). These base rates were obtained from estimates reported from a survey of the American Board of Clinical Neuropsychology members, so the accuracy of these estimates
have not been confirmed. Unfortunately, this study did not differentiate the concepts of malingering from symptom exaggeration, and the authors did not differentiate the base rates for FM from chronic fatigue syndrome. Higher rates of probable malingering were reported for cases referred by defense counsel and insurers in civil matters, and in criminal cases referred by prosecutors. The authors concluded that these base rates and diagnostic impressions might be partially influenced by the selection of cases by the referral source.

In a review of the literature, Weintraub (1995) cites studies showing that 20 to 46% of people consider deliberate misrepresentation of compensation claims to be acceptable behaviour. These studies suggest that exaggeration and malingering may not be rare, especially in situations where individuals are seeking compensation. This research supports the insurers’ vigilant approach to fraudulent claims. In fact, Malleson (2002) described FM as a medical entity that serves to maintain business for rheumatologists, stating that “...without these discontented fibromyalgics, the rheumatologists’ waiting rooms would be empty.” (p.168). He further states that “..rather than curing diseases, medicine is manufacturing new ones” (p.172).

Ross (1970) took a rare glimpse into claims that are settled out of court, and interviewed a number of claims adjusters in this process. He described the adjuster as one who believes in the validity of the formal law and endorses its assumptions concerning fault. Ross identified the adjuster’s central belief that claimants will distort evidence and exaggerate their losses, and that the adjuster is responsible to counteract the distortion and exaggeration before considering payment of the claim. Ross also found that claimants’ understanding of their rights and obligations are generally vague and imprecise. The claimant is often uninformed in evaluation and unskilled in negotiation, as opposed to the adjuster who has been trained in these skills.

According to Staples (1997), another approach used by the insurance industry to reduce costs is to evaluate whether claimants have taken steps to mitigate their problems or to avoid illness or injury. A life or health insurance relationship induces one into a contractual association with fitness and wellness practices, so that once a claim is made, the insurer reviews whether the claimant has been practicing behaviours to reduce loss. From this perspective, every claimant may be suspected of not doing enough to reduce loss, which may result in policy changes, such as higher premiums, exclusion clauses, and possibly denial of
future coverage. Ericson et al. (2000) has suggested that this approach to reduce industry costs can result in the claimant being viewed as the ‘offender’, as opposed to the ‘victim’. From this perspective, if a person slipped and injured themselves on a floor that had recently sustained a spill, that person may be suspect of increasing their risk of an injury by failing to notice and avoid the spill area, as opposed to not being responsible for the injury. According to Ericson, the insurer would suspect the claimant of at least partial responsibility for being injured, and a penalty to the claim would result. This point of view can lead to an assumption that may not be realistic if it is extended to many chronic medical conditions. In fact, it has been argued that treatment recommendations related to chronic pain conditions have been influenced by cost reduction strategies (Mersky & Teasell, 2000).

Of course, the insurer’s perspective must also be considered. For example, a worker who continually disregards safety guidelines and continues to engage in risky behaviour may eventually get injured. In this situation, the worker could have avoided the injury, and an increase in claimant costs would be a reasonable approach for the insurer to take under these circumstances.

The relationship between the insured and the insurer has been described as one involving “moral hazard”, whereby the insured’s incentive to avoid risky behaviour or practice behaviours to minimize risk is believed to be reduced by the ‘protection’ provided by insurance (Baker, 1996; Heimer, 1985). However, Ericson et al. (2000) characterize the insurance relationship as one that involves moral hazard for both the insured and the insurer, resulting in an inherent conflict in that relationship. In disability claims, it is in the interest of the insurer to recognize and cover as few medical problems as possible while convincing the consumer to purchase coverage. On the other hand, it is in the interest of the insured to have their medical problem recognized and to receive assistance and support. The increasing pressure to keep premiums down can lead the insurer to view claims as fraudulent, question claimant credibility, or otherwise be non-supportive of the claim, further leading to increased policing, surveillance, and denial of coverage. Diagnostic issues can play a significant role in the claims determination process, and chronic and controversial illnesses can be a target for raising credibility issues.

The issue of moral hazard has also been raised in doctor-patient relationships, in the context of disability assessments (The American Medical Association, 2000). Patients may
be influenced by monetary incentives to exaggerate their pain reports, thereby increasing their chance of obtaining disability coverage. Medical professionals performing disability evaluations have often been hired by insurance and government agencies hence, incentives may be present to doubt patient complaints. Merskey and Teasell (2000) have argued that there is a serious risk for treatment recommendations to be determined more by the interests of the insurer than by the actual needs of the patient, resulting in a disparagement of pain and disability. The conflicting interests of the insurer and insured, or doctor and patient in the context of a disability evaluation, continue to foster the bi-directionality of moral hazard.

**Disability Insurance Law**

In a legal context, more visible injuries and losses create an immediate expectation of impairment and a wish to compensate on the part of the judge or jury, whereas, the opposite expectation frequently occurs when there is no visible injury and medical evidence indicates no abnormalities (Oatley, 1997). Obviously and rightfully, judges, juries, and insurance carriers should be concerned about providing compensation for problems and injuries that appear to have no medical basis. That is precisely the situation in some claimants with psychological, psychiatric, or chronic pain conditions. The absence of compelling evidence of the problem, and disability associated with the problem, is not necessarily evidence of absence. Nonetheless, the onus of the burden of proof of impairment and disability is on the claimant; providing this evidence in cases involving non-obvious or controversial medical conditions can be challenging.

Disability claims have traditionally focused on claimants’ ability to provide an objective basis for their allegations of disability. However, by its very nature, FM cannot be substantiated on an objective basis (Carette, 1996). As described earlier, a diagnosis of FM relies on subjective criteria. Confusion surrounding the medical basis of FM is present in judicial disability determinations. The courts have not reached a consensus regarding the validity of FM, even though the disorder has been documented in the legal system since the 1950’s (Bierman, 1998). As such, plaintiff credibility is an increasingly important issue in the context of disability claims (Wolfe & Potter, 1996), especially because there is no simple method of assessing disability in FM claimants (Littlejohn, 1995).

According to Carmichael and Garson (1999), a disability policy typically describes the insured as either being incapable of carrying out the substantial or important duties of
their “own occupation” or being incapable of working at “any occupation” for which they are suited, based on their education, training, or experience. The “own occupation” definition commonly applies to the first period of disability (i.e. first two years) and the “any occupation” definition is applicable thereafter.

**Proof of Disability**

The basic test for proof of total disability is set out in the renowned passage from *Paul Revere Life Assurance v. Sucharov (1983)*:

The test of total disability is satisfied when the circumstances are such that a reasonable Man would recognise that he should not engage in certain activity even though he literally is not physically unable to do so. In other words, total disability does not mean absolute physical inability to transact any kind of business pertaining to one’s occupation, but rather that there is a total disability if the insured’s injuries are such that common care and prudence require him to desist from his business or occupation in order to effectuate a cure; hence, if the condition of the insured is such that in order to effect a cure or prolongation of life, common care and prudence will require that he cease work, he is totally disabled within the meaning of health or accident insurance policies.

The reference to a “reasonable man” is taken to impose an objective standard. The application of “reasonable” was dealt with in *Maslen v. Rubenstein (1994)*, wherein the court stated a need for “convincing” evidence that pain would continue beyond the “normal” recovery period, indicating that the plaintiff’s own evidence could suffice if it were consistent with the surrounding circumstances. However, objective evidence remains the most convincing form of evidence.

In the *Maslen* case, the court also attempted to provide some guidelines on dealing with cases that do not involve readily identifiable pathology, but more complex cases involving chronic pain conditions. In giving the judgment set out by the court, Taylor J. A. emphasized the importance of proceeding cautiously when the objective evidence of continuing injury is either scarce or absent. The judge indicated that the court must decide on
compensability based on issues of credibility and the explanations for the plaintiff’s condition. These guidelines have been extended in *Yoshikawa v. Yu (1994)*.

Diverse opinions have been expressed over controversial claims such as those involving FM. For example, in *Eddie v. Unum Life Insurance Company of Canada (1998)* the trial judge accepted the insured’s contention that she was disabled, and commented that disability coverage applied to someone who, based on presenting symptoms, perceived and felt they were unable to work. The case went to appeal and in the majority judgment, Prowse J. A. accepted the plaintiff’s subjective evidence; while a dissenting judge (Newbury J. A.) found that the plaintiff failed to prove the “cause” of her disability – fibromyalgia. In *Mathers v. Sun Life Assurance Co. (1998)* the insured complained of disability due to low back pain. The medical evidence was inconclusive. Even though the trial judge concluded that with no objective measurable evidence of disability, a compensable total disability due to subjective pain could be established, this particular claim was not supported due to insufficient medical evidence.

Carmichael and Garson (1999) point out that according to the current state of the law, no particular type of evidence, lay or expert, is required in order to prove disability. Objective medical evidence typically strengthens the case, but a claim can be found on more subjective evidence, especially if corroborated by lay or expert witnesses. For example, in *McCulloch v. City of Calgary (1985; pp. 237-238)* the insured complained of chest pain resulting from thoracic surgery. The insured was found to be totally disabled and the Court indicated an indifference to the cause of the pain, pointing out that the pain is a reality to the plaintiff.

Subjective evidence and credibility are addressed in a case involving FM in *Jones v. Prudential Group Assurance Co. (1999)*, where the Court made the statement reprinted below.

> During the trial we have heard orally from a number of medical experts, as well as the medical views filed with the court. These all deal with Fibromyalgia, as a syndrome in which the visual observations do not identify the problem. The complaints are not examinable, and the term, therefore, used, is that the complaints are subjective in nature, but this does not mean a disease does not exist. It
is only that science concerning the human body with all the advances made remains still imperfect as to the causes or basis for many of the human complaints with this disease. Fibromyalgia is classified as a syndrome, because science has not yet perfected an objective diagnosis for this disease.

The Court’s comment is illustrated in a historical account of medical views of multiple sclerosis. Some patients with multiple sclerosis experience severe pain in their bones, muscles, and skin, but many physicians questioned their reports of pain (Aring, 1973), until research in the early 1990’s confirmed that the disease could cause pain (see Ehde et al., 2003 for a review).

In establishing entitlement of benefits, the burden of proof is the usual civil standard of “the balance of probabilities” (Carmichael & Garson, 1999). The party asserting a state of affairs is usually the one that must provide proof. However, the court must decide who had the burden when the evidence is inconclusive. The issue in disability cases is the presence of disability, not the diagnosis of the underlying medical condition. A clearly identified diagnosis does not in itself prove disability. In the previously mentioned Eddie case, the insurer defended the case on the basis that FM was merely a label and not a condition recognized by medical science. However, the court commented that the fact of sickness was more important than its explanation. On appeal, the court upheld the original decision, stating that the insurer’s position was rejected because otherwise “situations could arise in which a claimant is clearly disabled by some kind of sickness, but is not eligible for benefits because the exact nature of the sickness cannot be determined.” (Prowse, J.A., para. 60). The judge’s comments clearly indicate that it is not essential to attach a diagnostic label or medical cause to an illness, rather, the importance is whether the person is legitimately disabled.

The diagnosis of an illness does not necessarily lead to disability. It is possible to suffer from some illness or impairment and not be disabled. For example, in the back pain case of Mathers (1998), the court noted that the finding of a chronic spasm was not “by itself evidence of a total disability caused by sickness or injury” (para. 48). The court further stated that “…there is a difference between pain and disability and a patient who complains of pain is not necessarily disabled by it.” (para. 52).
The American Medical Association provides guidelines to the evaluation of permanent impairment (2001). In an overview of those guidelines it states: (a) pain is subjective and does not lend itself to be readily validated or objectively measured, (b) the evaluation cannot be made on the basis of tissue pathology, (c) the evaluation must take into account biological, psychological, and social factors, and (d) physicians need to consider individual differences in their clinical judgment. Although these guidelines are not exclusively designed for court purposes, they are meant to assist physicians in the United States in the disability evaluation process. These guidelines clearly acknowledge that biological factors are only one facet in the assessment of chronic pain, regardless of the associated medical condition, a sentiment that has been echoed in several of the case judgments summarized earlier. The legal system will likely continue to struggle over the disability issue in FM even if objective evidence were available, as an objective finding is only one factor to be considered in resolving the question of disability.

**Expert Evidence**

An issue raised in cases of more subjective medical conditions is the weight of evidence given by doctors’ assessments. In *Maslen v. Rubenstein (1994)*, the judge commented that the court cannot take the doctors’ views of plaintiff reliability and truthfulness as deciding factors, but that the court itself must decide on critical issues of credibility and the balance of probabilities. The court similarly concluded that an expert’s testimony was inadmissible in *Kuhne v. Minifie (2000)*, because the expert’s opinions were related to secondary gain factors and plaintiff credibility, issues that were not within the domain of the expert. The role of the physician is considered to be one of diagnosis with possible comments on an individual’s level of impairment, whereas, the role of the court is to determine plaintiff credibility and whether the impairment results in disability, as defined in the insurance policy. This view has been similarly expressed by others, who indicate that proof of fraud is not the role of the health care professional (Bogduk, 2004; Mendelson, 1987; Mendelson & Mendelson, 2004). *Regina v. Marquard (1993)* was the leading case in which the judge confirmed the principle that the ultimate conclusion as to the truthfulness of a witness is for the trier of fact, and not within the scope of an expert witness. However, expert evidence of plaintiff credibility was deemed admissible in relation to the knowledge of psychological and physical factors that are considered beyond the experience of the lay
person. In this case, the judge accepted expert testimony related to plaintiff credibility from physicians who demonstrated professional experience with child abuse victims. Expert evidence of plaintiff credibility has also been accepted in association with results from validity scales and tests from psychological and physical assessments (e.g., *Calvez v. Illng*, 1997; *Lawson v. McGill*, 2004).

The credibility of expert evidence has been questioned in terms of secondary gain factors. For example, in *Calvez v. Illng*, (1997) the objectivity of an expert's opinion was raised, because she had received several referrals from the insurers of the defendant, her testimony was being submitted on behalf of the defendant, and of course, her fees were being paid for by the defendant. The issue of secondary gains biasing an expert's testimony can raise serious questions about the credibility of that expert's opinion, and influence the outcome of litigation.

The nature of proof in a disability claim will depend on the nature of the complaints, for example the proof of a broken limb is quite different from proof of a more subjective condition, such as depression. Wendell (1996) argues that medical professionals possess a level of authority within the courts, and a claimant's description of their physical condition is treated as weak evidence at best against medical descriptions. The result is that subjective descriptions of an individual's condition need confirmation to be accepted as accurate and credible. The best evidence of disability, from either the insurer's or the insured's perspective remains objective evidence, such as a physical or occupational capacity evaluation.

Scientific and medical evidentiary matters have been addressed in the courts, initially through *Frye v. United States* (1923). These matters have been addressed more recently and in greater detail in the influential case of *Daubert v. Merrell Dow Pharmaceuticals* (1993). In this case, the United States Supreme Court held that when expert evidence based upon "scientific knowledge" is offered at trial, the judge should act as a gatekeeper and first determine whether the proffered evidence is "reliable", and can be trusted to be scientifically valid. The specific factors explicated by the Daubert Court are: (1) whether the technique or theory can or has been tested, or objectively challenged; (2) whether the technique has been subjected to peer review and publication; (3) the known potential error rates; (4) the existence and maintenance of standards and controls; and (5) whether the technique has general acceptance in the scientific community. Similar criteria have been suggested in a
Credibility and Fibromyalgia

number of Canadian courts (e.g., *Grant v. Dube, 1992; Regina v. Mohan, 1994*). However, due to the subjective nature of phenomenon such as FM and chronic pain, there often is limited evidence that would meet the Daubert criteria in a disability claims situation.

**Mitigation, Contributory Negligence, Aggravation, and Punitive Damages**

Issues of mitigation, aggravation, and punitive damages may also play a role in a disability insurance claim. In the law of damages, mitigation refers to the conduct of the plaintiff that may diminish, or events that have diminished the loss complained of in the claim (Waddams, 1997). The basis of the rule of mitigation is that the plaintiff is not entitled to recover compensation for loss that could have been avoided, if reasonable action had been taken to avoid that loss. Therefore, failure to mitigate damages is a defense used by the insurer to reduce the plaintiff's claim. In issues of mitigation, tort and contract law generally accepts that a party has a “duty” to mitigate damages. For example, in personal injury cases, the plaintiff is expected to seek medical advice and to comply with reasonable medical treatment. In addition, the cost of the plaintiff's reasonable medical care is recoverable even if the care proves ineffective. The onus of proof is on the insurer to establish a failure to mitigate and demonstrate that steps in mitigation would be both realistic and beneficial to the insured.

Contributory negligence is also a defense used by the insurer, and refers to the plaintiff's conduct as it contributes to a loss in their claim. The rationale of this defence is that the law should assist only those persons considered worthy of protection, and individuals who consent to take on the risk of injury are denied the protection of the law (Linden, 2001). For example, if a person wore shoes with poor grips for his daily walk, and slipped and injured himself, that person would have contributed to the risk of injury in his selection of walking shoes. The onus of proof in this defense is again on the insurer to establish that the damage is a direct result of the risk assumed by the plaintiff.

Issues of aggravation generally refer to some form of “mental distress” and are a form of compensatory damages to the insured for the injury. For example, as previously mentioned in *Eddie v. Unum Life Insurance Company of Canada* (1998), it was held that the insurer wrongly demanded objective clinical evidence that was not required by the policy. As a result of the denial of benefits, the insured lost her home and suffered mental distress.
Punitive damages are non-compensatory by contrast, and are a form of punishment to the insurer for acrimonious, spiteful, malicious conduct. Punitive damages can be awarded when the insurer’s conduct clearly demonstrates a reprehensible breach of the duty of good faith, and are not linked to “mental suffering”. An additional requirement for the award of punitive damages is the element of a separate actionable wrong. It is a spectrum of misbehavior that continues to present a struggle in the courts, due to its apparent overlap with issues of aggravation. Punitive damages may be awarded when the insurer demands objective clinical evidence, knowing it is not required by the policy, and benefits are denied without allowing the insured time to provide evidence. Although punitive damages are extremely rare, they were awarded to a claimant with FM and chronic fatigue syndrome in a recent court of appeals case (*Fidler v. Sun Life Assurance Company, 2004*).

Consequently, the role played by each party is implied in the type of awards issued and damages declared in a disability claim. For example, the role and responsibility of the insured individual is central in claims involving mitigation, whereas, the insurer’s role becomes more central in claims involving aggravation and punitive damages. It is clear that controversy related to disability associated with FM is an ongoing and complicated process in the courts.

**Material Misrepresentation and Fraud**

According to the Insurance Act, R.S.B.C. 1996, c. 226, there are specific duties imposed on the involved parties. Those duties require that each party display *uberrima fides* or utmost good faith in their dealings. While it is not a fiduciary relationship, the relationship between insurer and insured is recognized as one where mutual obligations of trust and good faith are paramount in a contractual setting. Although truthfulness is also central in a tort setting, the duty of utmost good faith and disclosure between the involved parties is not imposed in a tort setting. The Insurance Act requires the insured in a contractual setting to truthfully disclose all requested facts, while the insurer is responsible for determining the relevance of those facts. False or incorrect answers, as well as missing information on a disability insurance application may put its continued existence in jeopardy (*Cavanaugh, 1999*). So, a policy of disability insurance may be voidable at the option of the insurer in the event of non-disclosure or misrepresentation of a material fact on the application, as has been the result in numerous cases (e.g., *Silva v. Sizoo, 1997*). In order for a claimant’s policy to be
beneficial, the claimant must be aware of and avoid issues arising from misrepresentation and non-disclosure on the insurance application.

During the initiation of the contract, the insured possesses knowledge of his personal affairs (e.g., income, health) and is required to truthfully disclose all material facts to the insurer. The challenge is that the insured does not know which facts are material to the contract, while the insurer knows what is material, but is unaware of each applicant’s facts. Any misrepresentation or non-disclosure must be “material” before it will have any effect on the validity of the insurance contract, and usually relates to the health or financial circumstances of the insured. The effects of a misrepresentation will vary, depending on the length of time after which the contract was formed. A material misrepresentation within the contestable period provides ground for voiding the contract regardless of whether it was made in innocence, even if the insured is unaware of the significance of the material facts.

The onus is on the insurer to prove that the three elements of fraud exist: (a) the applicant knew that the statement was false or reckless, (b) the false statement was material to the insurer, and (c) the false statement was calculated to mislead the insurer. For issues related to fraud, the insurer essentially must prove an intention to deceive. Proof of fraud requires a subjective test, which causes difficulty for insurers attempting to defend claims. Cases involving proof of insurance fraud typically illustrate a wide divergence in approaches taken by judges in dealing with issues of fraud and credibility. For example, in *Gregory v. Jolley* (1999), the court found the insured to be “generally truthful”, even though he was “not truthful” in the court testimony relating to his income. In a case comment found following an Alberta Court of Appeal case (*Milward v. Maritime Life Assurance Co.*, 1989), James A. Rendall indicated that the judgement was influenced by the court’s “serious misgivings about the claimant’s ‘insincerity’.”
Summary

FM is a very problematic condition due mainly to the substantial impact it can have on an individual’s physical functioning, disability level, and quality of life. As mentioned earlier, the number of claims being made against disability insurance companies has been increasing (Cavanaugh, 1999). Research suggests that exaggeration and malingering may not be uncommon, and it has already been demonstrated that conditions such as FM can successfully be feigned (e.g., Khostanteen et al., 2000). One review of the studies examining malingering and dissimulation suggests that malingering was present in 1.25 – 10.4% of chronic pain patients (Fishbain et al., 1999), while a second review study reported a base rate for malingering and symptom exaggeration of 35% for a combination of FM and chronic fatigue syndrome patients (Mittenberg et al., 2002). These findings suggest the need to search for more objective forms of evidence, and support the vigilant approach taken by insurers, in order to prevent an increase in their costs. However, Ericson et al. (2000) caution that there is an increasing pressure in the industry to keep premiums down, which might lower the threshold for the insurer to view claims as fraudulent. The interest of the insurer and insured fosters a bi-directional situation of moral hazard, wherein the relationship sets forth an inherent conflict of perspective between the insurer and the insured. The claims adjudication process is further complicated by the fact that FM criteria are based on self-reports and there is no definitive way to determine whether someone is disabled from FM. The controversy continues in the insurance industry as companies struggle to determine whether to cover disability claims for undetermined conditions such as FM, and patients with FM attempt to obtain coverage.

From a legal perspective, a disability claim for FM becomes especially complicated when the focus has traditionally been on the claimant’s ability to provide an objective basis for their allegations of disability. Even though the current state of the law indicates that no particular type of evidence is required in order to prove disability, the best evidence of disability, from either the insurer’s or the insured’s perspective remains objective evidence. There typically is scant evidence that would meet the Daubert criteria, or similar criteria suggested in the Canadian courts (e.g., Grant v. Dube, 1992; Regina v. Mohan, 1994), to validate claims made for certain medical conditions such as FM. The situation becomes especially challenging when the medical community continues to debate the defining criteria.
of FM. Confusion surrounding the medical basis of FM is equally profound in the context of judicial disability determinations. Therefore, it is important to explore and better understand procedures and methods of assessing credibility.

The research thus far supports the view that FM is multifaceted, with an interplay of biological and psychological factors. It may be that a disability evaluation can only be made in terms of the biopsychosocial model of disease and stated in terms of reasonable probabilities (Bennett, 1996). Due to its complex nature, the legal, medical, and insurance industries will continue to struggle with questions of credibility. Research in multiple domains is needed to fill in the missing pieces of this profoundly puzzling and controversial disorder.

Therefore, Study 1 is an examination of multiple domains involved in the disability determination process of patients with FM. Study 1 examines issues faced by patients with FM, the medical system, as well as an examination of issues faced by the legal system and insurance industry while they are involved in disability determinations.
Hypotheses

The case review study is focused on issues related to credibility in litigated cases involving FM in the Canadian courts. The study was exploratory in nature. A large number of demographic and credibility variables were collected. A review of the FM literature has led to two specific hypotheses for each of the three main factors under investigation, for a total of six hypotheses.

A significant majority of patients with a diagnosis of FM are women. The literature suggests that it is relatively common for credibility issues to be raised for women with health complaints, especially those with unclear etiologies (e.g., Ramsay, 1986; Register, 1987; Thorne, 1993). Therefore, the first two hypotheses are listed below.

1) The majority of litigants will be women (greater than 80%).

2) Women will be considered less credible (as determined by credibility ratings) and granted significantly lower total awards than men with FM.

The insurance industry literature suggests that issues related to industry costs, fraud, and malingering plays a significant role in the disability claims process. Chronic and controversial illnesses that rely predominantly on self-report can be a target for plaintiff credibility issues and add to the complexity to this process. The third and fourth hypotheses address these factors, as listed below.

3) Issues of misrepresentation, fraud, non-disclosure, failure to mitigate, and contributory negligence will be raised significantly more often in claims involving insurers than other defendants.

4) Significantly greater awards will be granted to plaintiffs in claims where investigative and surveillance data are not involved, than in claims where such information is involved.

Case judgments involving FM claims and similar chronic pain conditions suggest that plaintiff credibility and expert testimony play a significant role in litigation. Wendell (1996) suggested that medical professionals possess a greater level of authority than is appropriate within the courts. This opinion has been echoed in some case judgments (e.g., Maslen v. Rubenstein, 1994). Therefore, the final two hypotheses explore the credibility of plaintiff and expert testimony, and the impact they have on the outcome of this process, as listed below.
5) The courts will consider expert testimony to be significantly more credible (as determined by credibility ratings) than plaintiff testimony.

6) Expert testimony will be considered equally credible by the court, regardless of the expert’s familiarity with the plaintiff. (Familiarity with the plaintiff is measured in terms of whether the expert is a “treating” clinician or a “forensic” expert conducting an assessment solely for court purposes).
Method

Quicklaw is a comprehensive legal database that contains case judgments of every trial-by-judge alone (without jury) litigated case throughout the Canadian judicial system. This database was searched using the key word “Fibromyalgia” to identify all case judgments using the term for the time Quicklaw began in 1986 up to February of 2003. Those judgments were screened and only those cases involving awards sought for FM at the initial trial were included in the final data pool. Judgments that did not involve actions for FM awards were excluded (e.g., appeals, motions, and applications). Judgments printed in French were not included (n = 7), and all of these cases were tried in Quebec. A total of 194 cases were identified through that screening process. The cases were examined in relation to three major factors: demographic, insurance, and credibility. The epidemiological research indicates a high prevalence rate of FM among women. Gender related credibility issues have been raised in the FM literature. Therefore, a number of variables related to gender and credibility composed the demographic factor. Regional affiliation and associated information was also examined, given the disproportionate number of cases litigated in British Columbia (n = 119). A number of additional variables related to the demographic factor were also examined, such as age, marital status, and educational background.

The second major factor of interest involved the role that the insurance industry plays in the litigation of disability claims. The prevalence rate of insurers as defendants was examined, as well as the prevalence and credibility of surveillance information. A comparison of the total financial award granted was made with cases involving investigative and surveillance information and cases not involving such information, as an indicator of the impact that such information has on plaintiff success. Certain types of issues indicate court assigned liability to the plaintiff (e.g., contributory negligence, failure to mitigate, fraud, and misrepresentation) while certain types of damages indicate insurer liability (e.g., punitive, non-pecuniary/general, loss of past income, and future income). Therefore, these issues and types of damages acknowledged by the court were examined as indicators of court assigned liability.

Canadian insurance laws suggest that the focus should be on symptoms and functioning rather than medical diagnoses in disability determinations. Nonetheless, previous case judgments have tended to indicate that some form of objective evidence is crucial to
obtaining a financial award. These judgments further indicate that expert testimony is crucial for determining whether a plaintiff is successful in obtaining an award. Hence, a number of variables related to credibility issues composed the third factor, including: type and number of experts represented in court; familiarity of experts with the plaintiff (as determined by treating versus forensic expert); credibility of experts; inclusion of historical variables of plaintiff (i.e., treatment, employment, prior claims, and criminal history); type of objective evidence of disability; and credibility of the plaintiff (see Appendix A for a comprehensive list).

Credibility of testimony was determined by statements from the judge indicating level of agreement with the witness. The coding scheme was defined as follows: complete agreement = 2, partial agreement/some skepticism = 1, complete disagreement = 0.

The coding scheme for familiarity of the experts with the plaintiff was determined by statements from the judge indicating whether the doctor was a treating physician or an independent expert. The coding scheme was defined as follows: treating expert = 1, forensic expert = 2, mixed = 3. The “mixed” category represents those occasions when more than one expert witness with the same expertise testified, and those experts differed in their familiarity with the plaintiff. Whether the expert testified for the plaintiff or the defense also was coded.

Reliability

To examine interrater reliability of the variable coding scheme, three raters, blind to the experimental hypotheses, were familiarized with the case judgments and trained to code each variable. A detailed description of each variable was reviewed with the raters, and several case judgments were coded during training. Raters were given five case judgments to code independently, which were reviewed at the next training session. This process was repeated over three sessions, for a total of approximately 15 case judgments completed by each rater during training. Following training, each rater completed one-third of the 194 case judgments.

To assess the reliability of the credibility and familiarity coding schemes, as well as the remaining variables, the three raters coded a subsample of 10 percent of case judgments coded by another rater. A Cohen’s (1960) Kappa coefficient of .93 was obtained for overall interrater reliability of all variables, excluding credibility and familiarity ratings. A separate
Kappa of .87 was obtained for all credibility ratings, and a Kappa of .96 was obtained for all familiarity ratings.

**Statistical Analyses**

Variables were extracted from secondary records and content coded in each case judgment. The sample size was predetermined by the number of case judgments meeting inclusion criteria for the study (N=194). However, the number of subjects (i.e., case judgments) in each group comparison depended on which variables were being compared.

The relations between and among variables were examined using parametric and nonparametric statistics. Levene’s test for equality of variances was used to test for homogeneity of variance. The Kolmogorov-Smirnov procedure was used to determine whether the variables had non-normal distributions. Although the t-test is a rather robust parametric test with regards to violations of general linear model assumptions (e.g., homogeneity of variance and normally distributed variables), nonparametric analyses (e.g., Mann-Whitney U Test and Kruskal-Wallis Test) were also conducted on all variables because non-normal distributions were present for the majority of variables. Tukey HSD tests were computed following ANOVAs for unplanned comparisons. The Wilcoxon Sign Rank test was a nonparametric pairwise test used when samples sizes were small and unequal.

Although the probability of Type I error increases when multiple statistical comparisons are conducted, findings with alphas below .05 are reported. Less stringent significance levels were used because of the exploratory nature of this study.

**Preliminary Considerations Regarding Statistical Analyses**

Parametric and nonparametric statistics are presented throughout the Results section. In general, given the nature of the data, there were many, often gross, violations of assumptions underlying general linear model statistics. Sample sizes were sometimes very small or grossly discrepant. Most variables were not normally distributed. There were frequent violations of the assumption of homogeneity of variance. These distributional problems with many individual variables greatly increased the likelihood of both Type I and Type II statistical errors. Therefore, the reader must exercise caution when interpreting these results. In addition, the percentages of cases having certain characteristics are reported throughout the Results section. These percentages frequently do not add up to 100% due to missing data.
Results

Sample Characteristics

The gender distribution of the sample was 84% women (n = 163) and 16% men (n = 31). Their average age was 42 years (n = 153, SD = 10.6, Range = 16 - 69 years), and their average education was 13.4 years (n = 91, SD = 3.5, Range = 5 - 24 years). The marital status of the plaintiffs was 53% married/common law, 6% separated, 9% divorced, 10% single, .5% widowed, and 22% unknown. The mean number of children was 1.9 for each plaintiff (n = 95, SD = 1.1, Range = 0 – 5 children).

The initial diagnosis of FM was made by a rheumatologist (42%), family doctor (20%), physiatrist (16%), internist (3%), orthopedic surgeon (2%), neurologist (2%), neurosurgeon (2%), sports medicine physician (1%), chiropractor (1%), or other practitioner (2%; unknown for 11% of cases). Rheumatologists provided secondary/confirmatory diagnoses in 40% of the cases, followed by physiatrists (7%), a family doctor (3%), an orthopedic surgeon (2%), a neurologist, neurosurgeon, internist, and other practitioner (1% each, respectively). A secondary/confirmatory diagnosis was in agreement with the initial diagnosis in 41% of the cases, in disagreement in 21% of the cases, and unknown in 38% of the cases.

The FM claim was preceded by an injury or illness in 89% of the cases, while it appeared to occur independently of an injury or illness in 11% of the cases. For FM claims that were secondary to an injury or illness, the precipitating event was reported to be a motor vehicle accident in 79% of cases, a work-related accident in 2% of cases, a non-vehicle or work-related accident in 6% of cases, or the result of an illness or surgery, an assault, or a chiropractic adjustment (1% each, respectively).

In terms of a history of pain complaints, 61% of plaintiffs reported pain prior to the current claim, while 31% of plaintiffs did not report pain prior to the current claim. Of those plaintiffs who reported pain prior to the current claim, 18% of those complaints were related to a prior disability claim, while 47% were not (data was missing for the remaining 35%).

Co-existing medical diagnoses were included in the claim for 44% of the cases, while the remaining cases did not involve additional documented medical diagnoses. Of those cases involving co-existing medical diagnoses, 12% of the cases involved diagnoses for temporomandibular disorder, followed by 10% for chronic pain syndrome, 7% for chronic
fatigue syndrome, 5% for neurological problems (e.g., head injuries), 4% for low back pain, 4% for tendinitis, 3% for myofascial pain syndrome, 3% for whiplash, 3% for polyarthralgia NOS, and 3% for irritable bowel syndrome. A variety of additional co-existing medical diagnoses were also included in the claims, but the remaining diagnostic categories have less than a 3% prevalence rate.

In terms of variables related to mental health, co-existing psychological or psychiatric diagnoses were included in the claim for 42% of the case judgments. Of those cases, 66% involved diagnoses for depression, 2% for anxiety, 17% for a combination of depression and anxiety, 2% for other diagnoses. In terms of mental health history, 14% of the plaintiffs had received a psychological diagnosis prior to the claim, while 75% had not received such a diagnosis prior to the claim.

Details regarding awards granted for the total sample are summarized in Table 1. As indicated earlier, the table reveals great variability in sample size and non-normal distributions for the majority of award type samples. There were frequent violations of the assumption of homogeneity of variance, and the distributional problems greatly increased the likelihood of both Type I and Type II statistical errors.

Table 1
Awards and damages granted for total sample

<table>
<thead>
<tr>
<th>Award Type</th>
<th>n</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Normality Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Income</td>
<td>119</td>
<td>41,324</td>
<td>22,102</td>
<td>54,576</td>
<td>3.58</td>
<td>19.07</td>
<td>.0001</td>
</tr>
<tr>
<td>Future Income</td>
<td>61</td>
<td>106,552</td>
<td>60,000</td>
<td>118,086</td>
<td>1.42</td>
<td>1.32</td>
<td>.0001</td>
</tr>
<tr>
<td>Earning Capacity</td>
<td>31</td>
<td>95,524</td>
<td>50,000</td>
<td>105,041</td>
<td>1.49</td>
<td>1.80</td>
<td>.010</td>
</tr>
<tr>
<td>Special Damages</td>
<td>98</td>
<td>5,037</td>
<td>2,949</td>
<td>7,027</td>
<td>3.46</td>
<td>15.76</td>
<td>.0001</td>
</tr>
<tr>
<td>Agg/Punitive</td>
<td>5</td>
<td>9,200</td>
<td>8,500</td>
<td>3,718</td>
<td>.943</td>
<td>1.50</td>
<td>.759</td>
</tr>
<tr>
<td>Future Care</td>
<td>70</td>
<td>26,551</td>
<td>10,818</td>
<td>47,100</td>
<td>5.08</td>
<td>33.26</td>
<td>.0001</td>
</tr>
<tr>
<td>Non-pec/General</td>
<td>169</td>
<td>42,532</td>
<td>35,000</td>
<td>29,500</td>
<td>1.34</td>
<td>2.45</td>
<td>.0001</td>
</tr>
<tr>
<td>Total Award</td>
<td>191</td>
<td>130,395</td>
<td>70,000</td>
<td>170,363</td>
<td>2.49</td>
<td>7.18</td>
<td>.0001</td>
</tr>
</tbody>
</table>

Note. The Kolmogorov-Smirnov test was used to test normality.

Details regarding number of claims and type of awards granted per region are summarized in Tables 2 and 3. There was a disproportionate number of cases litigated in British Columbia (61%). The breakdown of cases by province was: British Columbia = 61%, Alberta = 11%, Saskatchewan = 5%, Manitoba = 1%, Ontario = 11%, New Brunswick = 5%, Nova Scotia = 4%, and the combined total of Newfoundland and Prince Edward Island = 2%.
Table 2
Awards & damages granted by geographic location: western Canada

<table>
<thead>
<tr>
<th>Award Type</th>
<th>British Columbia</th>
<th>Alberta</th>
<th>Saskatchewan</th>
<th>Manitoba</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Median (n)</td>
<td>Mean (SD)</td>
<td>Median (n)</td>
</tr>
<tr>
<td>Past Income</td>
<td>32,907 (35,720)</td>
<td>20,872 (79)</td>
<td>63,294 (72,417)</td>
<td>27,460 (10)</td>
</tr>
<tr>
<td>Future Income</td>
<td>81,817 (98,544)</td>
<td>49,000 (36)</td>
<td>76,410 (120,166)</td>
<td>18,592 (6)</td>
</tr>
<tr>
<td>Earning Capacity</td>
<td>94,589 (107,496)</td>
<td>52,500 (26)</td>
<td>198,333 (149,482)</td>
<td>198,333 (1)</td>
</tr>
<tr>
<td>Special Damages</td>
<td>3,591 (4,299)</td>
<td>2,310 (67)</td>
<td>7,658 (4,311)</td>
<td>7,658 (12)</td>
</tr>
<tr>
<td>Agg / Punitive</td>
<td>9,625 (4,151)</td>
<td>9,250 (4)</td>
<td>9,250 (1)</td>
<td>9,250 (1)</td>
</tr>
<tr>
<td>Future Care</td>
<td>16,619 (21,809)</td>
<td>8,193 (41)</td>
<td>75,513 (136,748)</td>
<td>14,721 (6)</td>
</tr>
<tr>
<td>Non-pec / General</td>
<td>43,183 (30,346)</td>
<td>35,000 (110)</td>
<td>40,617 (34,236)</td>
<td>34,187 (15)</td>
</tr>
<tr>
<td>Total Award</td>
<td>115,787 (141,331)</td>
<td>67,393 (119)</td>
<td>126,608 (228,039)</td>
<td>49,489 (20)</td>
</tr>
</tbody>
</table>

Note. The number of cases (N) located at the top of each column is only representative of the “Total Award” category. Abbreviations: Agg = aggravated; Non-pec = non-pecuniary. *Punitive damages were awarded only in Alberta, all remaining awards in that category are for aggravated damages.

Table 3
Awards & damages granted by geographic location: eastern Canada

<table>
<thead>
<tr>
<th>Award Type</th>
<th>Ontario</th>
<th>New Brunswick</th>
<th>Nova Scotia</th>
<th>NFL &amp; PEI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Median (n)</td>
<td>Mean (SD)</td>
<td>Median (n)</td>
</tr>
<tr>
<td>Past Income</td>
<td>71,198 (54,177)</td>
<td>70,333 (8)</td>
<td>42,711 (45,773)</td>
<td>27,500 (7)</td>
</tr>
<tr>
<td>Future Income</td>
<td>201,902 (151,601)</td>
<td>182,000 (9)</td>
<td>174,716 (129,682)</td>
<td>162,470 (5)</td>
</tr>
<tr>
<td>Earning Capacity</td>
<td>4,825 (12,62)</td>
<td>3,810 (6)</td>
<td>3,611 (4,726)</td>
<td>2,338 (5)</td>
</tr>
<tr>
<td>Special Damages</td>
<td>9,819 (12,162)</td>
<td>3,801 (6)</td>
<td>3,611 (4,726)</td>
<td>2,338 (5)</td>
</tr>
<tr>
<td>Agg / Punitive</td>
<td>67,768 (17,198)</td>
<td>66,000 (4)</td>
<td>12,416 (12,189)</td>
<td>8,700 (9)</td>
</tr>
<tr>
<td>Future Care</td>
<td>52,333 (28,777)</td>
<td>40,000 (15)</td>
<td>29,250 (10,412)</td>
<td>27,500 (10)</td>
</tr>
<tr>
<td>Non-pec / General</td>
<td>67,768 (28,777)</td>
<td>66,000 (15)</td>
<td>12,416 (10,412)</td>
<td>8,700 (10)</td>
</tr>
<tr>
<td>Total Award</td>
<td>192,384 (228,580)</td>
<td>56,367 (21)</td>
<td>175,585 (165,825)</td>
<td>142,574 (10)</td>
</tr>
</tbody>
</table>

Note. The number of cases (N) located at the top of each column is only representative of the “Total Award” category. Abbreviations: Agg = aggravated; Non-pec = non-pecuniary. *Punitive damages were awarded only in Alberta, all remaining awards in that category are for aggravated damages.
To examine differences between various types of awards granted and trial location, comparisons included total award, past income, future income, earning capacity, cost of future care, and non-pecuniary/general damages. Provinces with 20 or more case judgments were included in these comparisons (British Columbia, Alberta, and Ontario). A one-way ANOVA examining total award by trial location revealed no significant main effect \[ F(2,159) = 1.87, p < .157 \]. Levene’s test was significant, indicating a violation of homogeneity of variance. Total award by trial location was also compared using Kruskal-Wallis test because there were gross violations of general linear model assumptions. Results revealed no significant differences (p < .266). A one-way ANOVA examining non-pecuniary/general damages by trial location revealed no significant main effect \[ F(2,139) = \ 0.686, p < .505 \], and similar results were obtained using the Kruskal-Wallis test (p < .287).

A one-way ANOVA examining future income awards by trial location revealed a significant main effect, \[ F(2,50) = 4.37, p < .018 \]. Levene’s test was not significant, indicating that homogeneity of variance was not violated (although normality assumptions were violated). Follow-up comparisons revealed significantly greater future income awards for Ontario compared to British Columbia (Tukey HSD, p < .02, Cohen’s d = 1.1, large effect size). The results for future income approached significance using the Kruskal-Wallis test (p < .063). A one-way ANOVA examining past income awards by trial location revealed a significant main effect, \[ F(2,94) = 4.81, p < .01 \]. However, Levene’s test was significant, indicating a violation of homogeneity of variance. The results for past income approached significance using the Kruskal-Wallis test (p < .075), indicating greater past income awards for Ontario compared to British Columbia (Mann-Whitney U = 170.00, p<.032, d = 1.02, large effect size). A one-way ANOVA examining awards for special damages by trial location revealed a significant main effect, \[ F(2,84) = 6.52, p < .002 \]. Levene’s test was significant. The results for special damages also were significant using the Kruskal-Wallis test (p < .002). Follow-up comparisons revealed significantly greater awards of special damages for Alberta than British Columbia (Mann-Whitney U = 146.00, p <.0001, d = .95, large effect size). Finally, a one-way ANOVA examining cost of future care awards by trial location yielded a significant main effect, \[ F(2,50) = 5.33, p < .008 \]. Again, Levene’s test was significant. The results for future care were also significant using the Kruskal-Wallis test (p < .005), revealing significantly greater future care awards for Ontario than British Columbia.
Credibility and Fibromyalgia

(Mann-Whitney U = 7.00, p < .003, d = 2.39, very large effect size). The results of these analyses and follow-up comparisons are summarized in Table 4.

A comparison examining earning capacity by trial location could not be conducted because British Columbia was the only province with 20 or more case judgments that included awards for earning capacity. As well, comparison examining aggravated and punitive damages could not be conducted because there was only one case judgement involving punitive damages, and four case judgments involving aggravated damages.

Table 4
Summary of awards compared across three provinces

<table>
<thead>
<tr>
<th>Award Type</th>
<th>British Columbia N = 119</th>
<th>Alberta N = 20</th>
<th>Ontario N = 22</th>
<th>Post Hoc Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Income</td>
<td>Mean (SD) 32,907 (35,720)</td>
<td>Median (n) 20,872 (79)</td>
<td>Mean (SD) 63,294 (72,417)</td>
<td>Median (n) 27,460 (10)</td>
</tr>
<tr>
<td>Future Income</td>
<td>Mean (SD) 81,817 (98,544)</td>
<td>Median (n) 49,000 (36)</td>
<td>Mean (SD) 76,410 (120,166)</td>
<td>Median (n) 18,592 (6)</td>
</tr>
<tr>
<td>Special Damages</td>
<td>Mean (SD) 3,591 (4,299)</td>
<td>Median (n) 2,310 (67)</td>
<td>Mean (SD) 7,658 (4,311)</td>
<td>Median (n) 6,946 (12)</td>
</tr>
<tr>
<td>Future Care</td>
<td>Mean (SD) 16,619 (21,809)</td>
<td>Median (n) 8,193 (41)</td>
<td>Mean (SD) 75,513 (136,748)</td>
<td>Median (n) 14,721 (6)</td>
</tr>
<tr>
<td>Non-pec /General</td>
<td>Mean (SD) 43,183 (30,346)</td>
<td>Median (n) 35,000 (110)</td>
<td>Mean (SD) 40,617 (34,236)</td>
<td>Median (n) 30,000 (15)</td>
</tr>
<tr>
<td>Total Award</td>
<td>Mean (SD) 115,787 (141,331)</td>
<td>Median (n) 67,393 (119)</td>
<td>Mean (SD) 126,608 (228,039)</td>
<td>Median (n) 49,489 (228,580)</td>
</tr>
</tbody>
</table>

Note. The number of cases (N) located at the top of each column is only representative of the “Total Award” category. Abbreviations: Agg = aggravated; Non-pec = non-pecuniary; O=Ontario, A=Alberta, BC=British Columbia. P values with (*) indicate Kruskal-Wallis test results, otherwise, p values are from ANOVA results.

To examine the presence of temporal patterns in the number of FM claims litigated in Canada, the number of cases were calculated for each province, over six-year intervals (see Table 5). The results revealed an increase in the number of cases litigated for Nova Scotia, a decrease for British Columbia and Saskatchewan, and no particular trend for the remaining provinces.
Table 5
Number of cases litigated over time by geographic location

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>British Columbia</td>
<td>5</td>
<td>81</td>
<td>33</td>
</tr>
<tr>
<td>Alberta</td>
<td>0</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td>0</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Ontario</td>
<td>0</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>New Brunswick</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Nova Scotia</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>NFL &amp; PEI</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Gender Related Factors

Descriptive statistics by gender are summarized in Table 6. There were many more female plaintiffs (n = 163) than male plaintiffs (n = 31). The first hypothesis for gender related factors stated that the majority of litigants would be women (greater than 80%). Chi square analyses revealed a significant difference for gender distribution \(\chi^2(1,193) = 89.81, p < .0001\], showing significantly more women (84%) than men (16%) litigants.

The second hypothesis for gender related factors stated that women would be considered less credible and granted significantly lower total awards than men with FM. Due to a small and unequal sample size, this hypothesis was examined with Mann-Whitney U tests. There was no significant difference in credibility ratings between women and men (U = 2089.00, p < .08). As well, there was no significant difference in total award granted between women and men (U = 2238.00, p < .52). Additional analyses were conducted comparing total award granted between women and men with credibility ratings of “1” (partial agreement) and no significant differences were found (U = 1044.00, p < .709). A similar comparison of those with credibility ratings of “0” (complete disagreement) or “2” (complete agreement) could not be examined separately due to small sample sizes.

Table 6
Summary of credibility rating and total awards granted by gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>n</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>31</td>
<td>1.06</td>
<td>1.00</td>
<td>.51</td>
<td>128,688</td>
<td>47,279</td>
<td>195,430</td>
</tr>
<tr>
<td>Female</td>
<td>163</td>
<td>1.25</td>
<td>1.00</td>
<td>.56</td>
<td>130,714</td>
<td>71,884</td>
<td>165,959</td>
</tr>
</tbody>
</table>

Note. There was a nonsignificant trend toward women being rated as more credible than men (p < .08).
Credibility ratings were converted to a binary rating system, whereby credibility ratings of 0 were coded as “not credible” and ratings of 1 or 2 were coded as “credible”. Thus, 0 (complete disagreement) = “not credible”, while 1 (partial agreement), or 2 (complete agreement) = “credible”. This conversion allowed for an estimate of the percentage of people in each group rated as “credible”. A chi square analysis of the binary credibility ratings by gender revealed no significant differences $[\chi^2(1,193) = .51, p < .48]$, showing that women (93.8%) and men (90.3%) were similarly rated as “credible”.

**Insurance Factors**

Liability issues were reported in only 13% of case judgments. The defendants involved in FM claims, as identified by the name provided on the case judgment, were as follows: independent individuals (63%), individuals and business (17%), insurers (including public health, auto, work, and private insurers; 17%), estate (3%), and one instance of a provincial government (1%). The type of defendant could only be identified by the name provided on the case judgment. As a result, it is not known whether insurers may have also been involved in cases in which only an independent individual was identified as the defendant. However, FM was secondary to either a motor vehicle accident or a work related injury for the majority of plaintiffs (80%), and these claims almost certainly involved an insurer. However, the identification of insurer involvement should be interpreted with caution because of the absence of reliable information.

The third hypothesis stated that issues of plaintiff liability, including misrepresentation, fraud, non-disclosure, failure to mitigate, and contributory negligence would be raised significantly more often in claims involving insurers as defendants compared to other types of defendants. Given problems with identifying insurers, this hypothesis could not reliably be evaluated.

To examine the potential role that plaintiff liability played on the outcome of FM claims, total awards were compared between cases involving all possible plaintiff liability issues (contributory negligence, failure to mitigate, and misrepresentation/fraud/ non-disclosure) and cases with no plaintiff liability issues (see Table 7). A Mann-Whitney U test revealed significantly greater total awards for cases with no plaintiff liability issues than cases involving plaintiff liability issues ($U = 1394.00, p < .008, d = .48$, medium effect size). Mann-Whitney U tests examining specific types of plaintiff liability issues revealed
significantly greater total awards only when cases with no liability issues were compared to cases involving issues of misrepresentation or non-disclosure ($U = 48.50, p<.001, d = .79$, large effect size), but cases involving contributory negligence ($U = 286.00, p<.142$), or failure to mitigate revealed no significant difference between groups ($U = 1217.50, p<.388$). The examination of separate liability issues involved extremely small sample sizes, and should therefore be interpreted with caution.

Table 7

<table>
<thead>
<tr>
<th>Issues of Liability</th>
<th>Issues Present</th>
<th>Issues Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (SD)</td>
<td>Mean Award (SD)</td>
</tr>
<tr>
<td>Contributory Negligence</td>
<td>5 (62,058)</td>
<td>48,244 (62,058)</td>
</tr>
<tr>
<td>Failure to Mitigate</td>
<td>16 (78,296)</td>
<td>80,463 (175,823)</td>
</tr>
<tr>
<td>Misrepresentation/ Fraud/ Non-disclosure*</td>
<td>5 (3,783)</td>
<td>2,680 (171,333)</td>
</tr>
<tr>
<td>Total Liability Issues</td>
<td>25 (72,785)</td>
<td>61,821 (178,411)</td>
</tr>
</tbody>
</table>

*None of the claims involved issues of fraud.

Descriptive statistics for plaintiff liability issues in relation to specific awards are provided in Table 8. Mann-Whitney U tests revealed significantly greater past income, future income, and future care awards for cases with no plaintiff liability issues than cases involving plaintiff liability issues. Mann-Whitney U test results approached significance for general damages, but not for earning capacity, or special damages. Again, some of these comparisons involve extremely small sample sizes, and should therefore be interpreted with caution.
The fourth hypothesis stated that significantly greater awards would be granted to plaintiffs in claims where investigative and surveillance data are not involved, than in claims where such information is involved. Surveillance information was considered present if it was reported in the case judgment and absent if it was not. The amount of award was not significantly related to the presence of investigative and surveillance information (Mann-Whitney U = 2543.00, \( p < .825 \)). Details regarding the number of cases involving surveillance information and amount of award granted are listed in Table 9. Of the cases in which surveillance was reported, 91% included videotaping, 6% included direct observation, and 3% included collateral information. The credibility of the surveillance information will be examined in the next section.

Table 9
Amount of award granted relative to presence of surveillance information

<table>
<thead>
<tr>
<th>Surveillance Information</th>
<th>n</th>
<th>Mean Award Granted</th>
<th>Median Award Granted</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>33</td>
<td>157,724</td>
<td>53,968</td>
<td>231,432</td>
</tr>
<tr>
<td>Absent</td>
<td>158</td>
<td>124,687</td>
<td>73,771</td>
<td>154,975</td>
</tr>
</tbody>
</table>
Credibility and Fibromyalgia

Credibility Factors

Descriptive statistics related to award granted by credibility ratings are presented in Table 10. To examine the role that plaintiff credibility played on the total award granted, a comparison of total award granted was examined between the judge’s credibility ratings of “0” (no agreement), “1” (partial agreement), and “2” (complete agreement) on plaintiff credibility. A one-way ANOVA examining total award granted by plaintiff credibility ratings revealed a significant main effect \[F(2, 187) = 7.67, p < .001\]. Levene’s test was significant, indicating a violation of homogeneity of variance. This violation may have been due to the small sample size of the plaintiff group receiving a credibility rating of “complete disagreement”. Plaintiffs with good credibility ratings (“complete agreement”) received much larger awards than plaintiffs with poor credibility ratings (“complete disagreement”; Tukey HSD, \(p < .001\), \(d = 1.1\), large effect size) or plaintiffs with medium credibility ratings (“partial agreement”; Tukey HSD, \(p < .013\), \(d = .45\), medium effect size). There was a trend toward plaintiffs with medium credibility ratings (“partial agreement”) receiving larger awards than plaintiffs with poor credibility ratings (“complete disagreement”; Tukey HSD, \(p < .077\), \(d = .74\), medium effect size). A Kruskal-Wallis test was also conducted, which yielded a significant difference in amount of award granted by plaintiff credibility ratings \(p < .0001\). Mann-Whitney U test comparisons similarly revealed that plaintiffs with good credibility ratings (“complete agreement”) received much larger awards than plaintiffs with poor (“complete disagreement”; \(U = 38.00\), \(p < .0001\), \(d = 1.08\)), or medium credibility ratings (“partial agreement”; \(U = 2358.00\), \(p < .002\), \(d = .45\)), and plaintiffs with medium credibility ratings (“partial agreement”) received much larger awards than plaintiffs with poor credibility ratings (“complete disagreement”; \(U = 182.50\), \(p < .0001\), \(d = .74\)).

Table 10

<table>
<thead>
<tr>
<th>Issues of Credibility*</th>
<th>Mean Award Granted if “2”</th>
<th>Median (SD)</th>
<th>Mean Award Granted if “1”</th>
<th>Median (SD)</th>
<th>Mean Award Granted if “0”</th>
<th>Median (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plaintiff</td>
<td>$189,981</td>
<td>123,938</td>
<td>$114,245</td>
<td>63,559</td>
<td>$10,613</td>
<td>5,400</td>
</tr>
<tr>
<td></td>
<td>(n = 55)</td>
<td>(201,778)</td>
<td>(n = 122)</td>
<td>(153,124)</td>
<td>(n = 13)</td>
<td>(14,456)</td>
</tr>
<tr>
<td>Surveillance**</td>
<td>$62,500</td>
<td>45,909</td>
<td>$79,431</td>
<td>40,750</td>
<td>$231,203</td>
<td>298,906</td>
</tr>
<tr>
<td></td>
<td>(n = 16)</td>
<td>(70,240)</td>
<td>(n = 8)</td>
<td>(96,838)</td>
<td>(n = 5)</td>
<td>(181,580)</td>
</tr>
</tbody>
</table>

* \(0\)=complete disagreement, \(1\)=partial agreement, \(2\)=complete agreement

** Surveillance evidence was reported in 17% of cases.
Descriptive statistics related to award granted by surveillance credibility ratings are presented in Table 10. Although surveillance evidence was reported in 17% of cases, awards were granted in 15% of cases involving surveillance. To examine the role that surveillance credibility played on total award granted, a comparison of total award granted was examined between the judge’s credibility ratings of "0", "1", and "2" on surveillance information. A one-way ANOVA examining total award granted by surveillance credibility ratings revealed a significant main effect \( F(2, 26) = 5.34, p < .011 \). Levene’s test was significant, and this violation may have been partially due to small sample sizes. Cases with good surveillance credibility ratings ("complete agreement") received much smaller awards than cases with poor surveillance credibility ratings ("complete disagreement"); Tukey HSD, \( p < .009 \), \( d = 1.74 \), very large effect size). Cases with medium surveillance credibility ratings ("partial agreement") approached significance in receiving smaller awards than cases with poor surveillance credibility ratings ("complete disagreement"; Tukey HSD, \( p < .039 \), \( d = 1.17 \), large effect size). Significant differences were not found in award amounts between cases with good ("complete agreement") and medium ("partial agreement") surveillance credibility ratings (Tukey HSD, \( p < .923 \), \( d = .21 \)). The Kruskal-Wallis test indicated no significant difference between surveillance credibility ratings and amount of award granted (\( p < .162 \)). This result may be due to the small sample sizes involved, and the conservative nature of this test.

To examine the potential role that liability issues played in relation to plaintiff credibility ratings, plaintiff credibility was examined in cases involving all possible liability issues (contributory negligence, failure to mitigate, and misrepresentation/fraud/ non-disclosure) and cases with no liability issues. A Mann-Whitney U test revealed significantly greater credibility ratings for plaintiffs with no liability issues than plaintiffs with liability issues (\( U = 1331.50, p < .001 \), \( d = .81 \)). Mann-Whitney U tests examining specific types of liability issues revealed significantly greater credibility ratings when cases with no liability issues were compared to cases involving issues of misrepresentation or non-disclosure (\( U = 88.50, p < .001 \), \( d = 1.99 \)), and approached significance for cases involving contributory negligence (\( U = 294.00, p < .09 \), but not for cases involving failure to mitigate (\( U = 1183.50, p < .197 \)). These results are presented in Table 11. The examination of separate liability issues involved extremely small sample sizes, and should therefore be interpreted with caution.
Table 11

Plaintiff credibility relative to the presence of liability issues

<table>
<thead>
<tr>
<th>Issues of Liability</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Issues Absent</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contributory Negligence</td>
<td>5</td>
<td>.80</td>
<td>.45</td>
<td>188</td>
<td>1.23</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>Failure to Mitigate</td>
<td>16</td>
<td>1.06</td>
<td>.44</td>
<td>177</td>
<td>1.24</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>Misrepresentation/Fraud/Non-disclosure</td>
<td>5</td>
<td>.20</td>
<td>.45</td>
<td>188</td>
<td>1.25</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td>Total Liability Issues</td>
<td>25</td>
<td>.84</td>
<td>.55</td>
<td>168</td>
<td>1.28</td>
<td>.54</td>
<td></td>
</tr>
</tbody>
</table>

The average credibility ratings for the plaintiff and for various experts are presented in Table 12. In general, the experts had higher credibility ratings than the plaintiff (scaled from 0 to 2). The fifth hypothesis stated that the courts would consider expert testimony to be significantly more credible than plaintiff testimony. Plaintiff credibility ratings were compared to expert credibility ratings using paired sample $t$-tests, and only experts with a group membership of 20 or more were reported. Pairwise comparisons revealed significantly greater credibility ratings for general practitioners, rheumatologists, physiatrists, and psychiatrists/psychologists testifying for the plaintiff ($p < .0001, .0001, .036, \text{ and } .003$ respectively), and for rheumatologists testifying for the defense ($p < .033$), compared to plaintiff credibility ratings.

Table 12

Plaintiff credibility relative to expert credibility ratings

<table>
<thead>
<tr>
<th>Pairwise Comparison</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>t score</th>
<th>p value</th>
<th>Effect Size (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plaintiff</td>
<td>83</td>
<td>1.23</td>
<td>.59</td>
<td>-4.045</td>
<td>.0001</td>
<td>0.57</td>
</tr>
<tr>
<td>Plaintiff GP</td>
<td>1.56</td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaintiff</td>
<td>84</td>
<td>1.20</td>
<td>.58</td>
<td>-4.599</td>
<td>.0001</td>
<td>0.65</td>
</tr>
<tr>
<td>Plaintiff Rheumatologist</td>
<td>1.56</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaintiff</td>
<td>47</td>
<td>1.28</td>
<td>.54</td>
<td>-2.162</td>
<td>.036</td>
<td>0.4</td>
</tr>
<tr>
<td>Plaintiff Physiatrist</td>
<td>1.52</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaintiff</td>
<td>38</td>
<td>1.21</td>
<td>.66</td>
<td>-1.673</td>
<td>.103</td>
<td>0.4</td>
</tr>
<tr>
<td>Plaintiff Orthopedic Surgeon</td>
<td>1.46</td>
<td>.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaintiff</td>
<td>44</td>
<td>1.09</td>
<td>.56</td>
<td>-3.092</td>
<td>.003</td>
<td>0.57</td>
</tr>
<tr>
<td>Plaintiff Psychiatrist/Psychologist</td>
<td>1.43</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaintiff</td>
<td>41</td>
<td>1.05</td>
<td>.59</td>
<td>-2.208</td>
<td>.033</td>
<td>0.5</td>
</tr>
<tr>
<td>Defense Rheumatologist</td>
<td>1.39</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaintiff</td>
<td>25</td>
<td>1.08</td>
<td>.57</td>
<td>-5.92</td>
<td>.559</td>
<td>0.17</td>
</tr>
<tr>
<td>Defense Orthopedic Surgeon</td>
<td>1.20</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Plaintiff sample sizes were equal to expert sample sizes for each comparison.
Plaintiff and expert credibility ratings were converted to a binary rating system, whereby credibility ratings of 0 were coded as “not credible” and ratings of 1 or 2 were coded as “credible”. Thus, 0 (complete disagreement) = “not credible”, while 1 (partial agreement), or 2 (complete agreement) = “credible”. This conversion allowed for an estimate of the percentage of people in each group rated as “credible” (see Table 13). To explore the converted rating system, binary credibility ratings of plaintiffs were compared to defense orthopedic surgeons. A chi square analysis revealed no significant differences in the percentage of plaintiffs and defense orthopedic surgeons considered “credible” \( \chi^2(1,24) = .16, p < .69 \). Given that defense orthopedic surgeons had the greatest difference of percentage rated as “credible” from plaintiffs, the results also suggest that the percentage of plaintiffs considered “credible” are similar to the percentages of experts considered “credible”.

Table 13
Percentage of plaintiffs and experts rated as credible

<table>
<thead>
<tr>
<th>Type of Witness</th>
<th>Number of Cases</th>
<th>Credible</th>
<th>Number of Cases</th>
<th>Not Credible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plaintiff</td>
<td>180</td>
<td>93%</td>
<td>13</td>
<td>7%</td>
</tr>
<tr>
<td>Plaintiff GP</td>
<td>80</td>
<td>96%</td>
<td>3</td>
<td>4%</td>
</tr>
<tr>
<td>Plaintiff Rheumatologist</td>
<td>82</td>
<td>98%</td>
<td>2</td>
<td>2%</td>
</tr>
<tr>
<td>Plaintiff Physiatrist</td>
<td>43</td>
<td>91%</td>
<td>4</td>
<td>9%</td>
</tr>
<tr>
<td>Plaintiff Orthopedic Surgeon</td>
<td>36</td>
<td>95%</td>
<td>2</td>
<td>5%</td>
</tr>
<tr>
<td>Plaintiff Psychiatrist/Psychologist</td>
<td>41</td>
<td>91%</td>
<td>4</td>
<td>9%</td>
</tr>
<tr>
<td>Defense Rheumatologist</td>
<td>34</td>
<td>83%</td>
<td>7</td>
<td>17%</td>
</tr>
<tr>
<td>Defense Orthopedic Surgeon</td>
<td>19</td>
<td>76%</td>
<td>6</td>
<td>24%</td>
</tr>
</tbody>
</table>

The final hypothesis stated that expert testimony would be considered equally credible by the court, regardless of the expert’s familiarity with the plaintiff. This hypothesis was examined by comparing treating clinicians with forensic (independent) experts. Given the minimal information found in the case judgments regarding expert familiarity, the results of these analyses should be interpreted with caution. Wilcoxon Signed Rank tests were used because they provide more sensitivity to smaller sample sizes than other paired tests, and sample size was small for many of these groups. Only experts with one group membership of 20 or more were reported. Wilcoxon Signed Rank tests revealed that credibility was not significantly related to familiarity with the plaintiff. These results are summarized in Table
14. As an additional analysis, credibility ratings for each of the expert categories in Table 14 were combined across treating and forensic categories to increase sample size, resulting in one treating and one forensic expert group. A paired sample $t$-test revealed significantly ($p < .05$) greater credibility ratings for treating clinicians (1.52), compared to forensic experts (1.25). The combined treating and forensic expert group credibility ratings were also converted to the binary credibility rating system described earlier. To explore the converted rating system, binary credibility ratings of treating clinicians were compared to forensic experts. A chi square analysis revealed no significant differences in the percentage of treating clinicians and forensic experts considered “credible” [$\chi^2(1,49) = .24, p < .63$].

Table 14

<table>
<thead>
<tr>
<th>Expert Witness</th>
<th>Familiarity Status</th>
<th>n</th>
<th>Median</th>
<th>Mean</th>
<th>SD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatologist</td>
<td>Treating</td>
<td>80</td>
<td>2.0</td>
<td>1.55</td>
<td>.53</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>Forensic</td>
<td>40</td>
<td>2.0</td>
<td>1.43</td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td>Physiatrist</td>
<td>Treating</td>
<td>43</td>
<td>2.0</td>
<td>1.55</td>
<td>.66</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>Forensic</td>
<td>10</td>
<td>1.0</td>
<td>1.00</td>
<td>.82</td>
<td></td>
</tr>
<tr>
<td>Orthopedic Surgeon</td>
<td>Treating</td>
<td>37</td>
<td>1.5</td>
<td>1.46</td>
<td>.59</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>Forensic</td>
<td>25</td>
<td>1.0</td>
<td>1.20</td>
<td>.82</td>
<td></td>
</tr>
<tr>
<td>Psychiatrist/Psychologist</td>
<td>Treating</td>
<td>43</td>
<td>1.7</td>
<td>1.44</td>
<td>.61</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>Forensic</td>
<td>16</td>
<td>1.0</td>
<td>1.05</td>
<td>.79</td>
<td></td>
</tr>
</tbody>
</table>

To further explore expert testimony, the credibility of experts testifying for the plaintiff was compared with experts testifying for the defense. Only experts with one group membership of 20 or more were reported. Wilcoxon Signed Rank tests revealed that credibility was not significantly related to testimony status. These results are summarized in Table 15. As a secondary analysis, credibility ratings for rheumatologists and orthopedic surgeons were combined across plaintiff and defense testimony categories to increase the sample size. A Wilcoxon Signed Rank test again revealed that credibility ratings were not significantly related to testimony status ($p < .325$).
<table>
<thead>
<tr>
<th>Expert Witness</th>
<th>Testimony</th>
<th>n</th>
<th>Median</th>
<th>Mean</th>
<th>SD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatologist</td>
<td>Plaintiff</td>
<td>84</td>
<td>2.0</td>
<td>1.56</td>
<td>.53</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>Defense</td>
<td>41</td>
<td>2.0</td>
<td>1.39</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>Physiatrist</td>
<td>Plaintiff</td>
<td>47</td>
<td>2.0</td>
<td>1.52</td>
<td>.66</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>Defense</td>
<td>9</td>
<td>1.0</td>
<td>1.0</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>Orthopedic Surgeon</td>
<td>Plaintiff</td>
<td>38</td>
<td>1.58</td>
<td>1.46</td>
<td>.59</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>Defense</td>
<td>25</td>
<td>1.0</td>
<td>1.2</td>
<td>.82</td>
<td></td>
</tr>
<tr>
<td>Neurologist</td>
<td>Plaintiff</td>
<td>20</td>
<td>2.0</td>
<td>1.68</td>
<td>.65</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Defense</td>
<td>7</td>
<td>2.0</td>
<td>1.57</td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td>Psychiatrist/Psychologist</td>
<td>Plaintiff</td>
<td>45</td>
<td>1.66</td>
<td>1.42</td>
<td>.63</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>Defense</td>
<td>17</td>
<td>1.0</td>
<td>1.05</td>
<td>.76</td>
<td></td>
</tr>
</tbody>
</table>
Discussion

Study 1 is unique in that it is a first step in the investigation of issues related to credibility in cases involving FM in the Canadian legal system. Hence, the purpose of the first study was more exploratory in nature, with the focus on examining differences and patterns in the variables under investigation. A large number of variables were examined in these cases, including variables related to geographic region, gender, insurance, and credibility issues.

In terms of diagnostic issues, a rheumatologist made the initial diagnosis of FM in more cases (42%) than other types of physicians, and provided a secondary/confirmatory diagnosis in more cases (40%) than other types of physicians. This finding is in keeping with the medical literature, which indicates that the diagnostic criteria for FM were formulated by the American College of Rheumatology (Wolfe et al., 1990), and considered to be a type of rheumatic disease. Co-existing medical diagnoses were included in 44% of the cases, and temporomandibular disorder (TMD; 12%) was the most frequent condition. This finding is consistent with the medical literature, which indicates that 13% to 18% of patients with FM have a comorbidity rate of TMD (Aaron, Burke, & Buchwald, 2000; Plesh, Wolfe, & Lane, 1996). A number of other medical diagnoses included in the claims, like chronic pain, fatigue, and irritable bowel syndrome, are also consistent with additional problems associated with FM in the medical literature (Baumstark & Buckelew, 1992; Smythe, 1989; Uveges et al., 1990; Wolfe et al., 1990). Some of the remaining co-existing diagnoses identified in the claims were more likely related to the type of injury preceding the onset of the FM. For example, neurological problems (e.g., head injuries), low back pain, and whiplash would not be unusual diagnoses for plaintiffs involved in motor vehicle accidents prior to the onset of FM.

Psychological or psychiatric diagnoses were included in the FM claim in 42% of the case judgments, and a number of those cases involved diagnoses of depression (66%), anxiety (2%), and a combination of depression and anxiety (17%). These findings are consistent with the research reporting high prevalence rates of depression and anxiety in patients with FM (Goldenberg, 1986; Hudson et al., 1992; Hudson et al., 1985; Walker et al., 1997a).
The FM claim was preceded by an injury or illness in 89% of the cases, while it appeared to occur independently of an injury or illness in 11% of the cases. These results are somewhat inconsistent with a previous study that examined the frequency of a precipitating event occurring prior to the onset of FM (Greenfield, Fitzcharles, & Esdaile, 1992). In that study, FM occurred independently of an injury or illness for 77% of the sample (primary), and FM was preceded by an injury or illness for 23% of the sample (reactive). Their patient sample was derived from a hospital outpatient rheumatology practice and a private rheumatology practice, with no indication of involvement in litigation, whereas the current sample was derived entirely from a sample of litigants. Clearly, the current sample is more likely to have FM that was preceded by an injury or illness, otherwise they would not be in litigation.

There were a disproportionate number of FM cases litigated in British Columbia (61%), compared to the remaining provinces. One possible explanation for this discrepancy is that some provinces have “no-fault” or “modified no-fault” insurance systems for personal injury claims (e.g., Saskatchewan, Manitoba, and Quebec). This type of system would make access to trial for FM claimants more difficult than in British Columbia, which has a pure tort system. In the “no-fault” system, the claimant is only allowed to access a legislated amount of compensation, and an appeals process is available. The data provides partial support for this explanation, given that the majority of FM claims (89%) were preceded by an injury or illness, thereby increasing the likelihood of a personal injury claim. A second possible explanation is the level of influence a public or private insurer may have on the amount of litigation. An insurer may be less likely to litigate and more prone to settling out of court if they are interested in reducing their costs. Therefore, the issue might be a matter of who is funding the litigation, and the status of their funding resources. The greater the funding resources, the more likely litigation becomes an option. A third possible explanation is that the history of the judiciary may differ in each province regarding the handling of FM claims. For example, the judiciary in Alberta displayed strong opposition to a FM claim (i.e., Mackie v. Wolfe, 1994), which may have sent a message to the provincial bench and bar regarding the handling of future claims. A final possible explanation is that there may be a higher prevalence of FM diagnoses in BC compared to other provinces, or that the diagnostic practices of rheumatologists may differ for FM across provinces.
In regards to the amount of awards grated by location, there were no significant differences between provinces for total award granted or non-pecuniary/general damages. However, there were significant differences between British Columbia and Ontario for some awards, with Ontario granting significantly greater awards for past and future income, and cost for future care.

The first hypothesis was related to gender distribution and stated that the majority of litigants would be women (greater than 80%). The gender distribution of the sample was 84% women and 16% men. The results confirm the first hypothesis, revealing significantly more women than men litigants. These findings are similar to the medical research, which indicates that FM is more common in women than men (Bennett et al., 1992; Wolfe et al., 1995). The second hypothesis was also related to gender and stated that women would be considered less credible (as determined by credibility ratings) and granted significantly lower total awards than men with FM. The results revealed no significant difference between women and men in credibility ratings or awards, and did not support the second hypothesis. In fact, there was a nonsignificant trend toward women being rated as more credible than men.

The third hypothesis stated that issues of plaintiff liability, including misrepresentation, fraud, non-disclosure, failure to mitigate, and contributory negligence would be raised significantly more often in claims involving insurers as defendants compared to other types of defendants. An examination of plaintiff liability issues revealed that they were reported in 13% of the case judgments. However, defendants were identified by the name provided on the case judgment. Therefore, it is not known whether insurers may have been involved in cases in which only an independent individual was identified as the defendant, as well as in cases where insurers were directly identified. Because FM was secondary to either a motor vehicle accident or a work related injury for the majority of plaintiffs (80%), these claims almost certainly involved an insurer. The amount of insurer involvement should be interpreted with caution because of the absence of reliable information. Given the problems with identifying insurers, the third hypothesis could not reliably be evaluated.

The role and responsibility of the plaintiff is central in claims involving issues of misrepresentation, fraud, non-disclosure, failure to mitigate, and contributory negligence.
The presence of these issues suggests a possible decrease or loss in the claim as a result of the plaintiff's conduct. Plaintiff liability issues were present in 13% of cases. The potential role of plaintiff liability was examined by comparing total award granted for cases involving plaintiff liability issues and cases with no plaintiff liability issues. The absence of plaintiff liability issues was associated with increased amounts of award granted, as would be expected. Specifically, greater past income, future income, and future care awards were given in cases with no plaintiff liability issues than cases involving plaintiff liability.

Misrepresentation and non-disclosure had an enormous impact on size of award, so that the total award granted was approximately 50 times more in cases where these issues were absent than when they were present.

The role and responsibility of the defendant is central in claims involving aggravated and punitive damages, and the presence of these damages suggest a possible increase in the claim as a result of the defendant's conduct. Aggravated and punitive damages were awarded in only 3% of the total cases, while plaintiff liability issues were present in 13% of cases. These results suggest that the plaintiff was responsible in more cases than the defendant in decreasing the award granted.

The fourth hypothesis stated that significantly greater awards would be granted to plaintiffs in claims where investigative and surveillance data are not involved, than in claims where such information is involved. Investigative and surveillance information was reported in only 17% of case judgments, and that information was examined in relation to award granted. The presence of such information did not have an impact on the amount of award, indicating no support for the fourth hypothesis. However, judges' perceptions of the credibility of the surveillance evidence were highly related to the amount of award granted. "Persuasive" or credible surveillance data calling into question plaintiff credibility resulted in awards that were 27% the size of the awards granted to cases in which the surveillance data was not considered persuasive.

Similarly, judges' perceptions of plaintiff credibility clearly were related to the amount of award granted. Those rated as most credible received awards 18 times greater than those rated as least credible. Although not statistically significant, there was a similar trend for plaintiffs with medium credibility ratings to receive larger awards than those rated least credible. Therefore, plaintiff credibility has a major impact on the outcome of a claim.
also important to note that the majority of plaintiffs (63%) received a medium credibility rating, suggesting that judges were somewhat uncertain about the credibility of most plaintiffs.

An examination of the effects of liability issues on plaintiff credibility revealed that judges perceived plaintiffs to be more credible when liability issues were not present for the plaintiff. Plaintiffs were perceived by judges to be much less credible in cases involving misrepresentation or non-disclosure than other plaintiff liability issues. These results suggest that a plaintiff's conduct can influence their credibility, as perceived by the courts.

The fifth hypothesis stated that the courts would consider expert testimony to be significantly more credible than plaintiff testimony. A comparison of plaintiffs' and experts' credibility indicated judges generally perceived the experts to be more credible than the plaintiffs. These findings support the fifth hypothesis, and may be specific to FM claims or other claims involving undetermined pain conditions, but may not be generalizable to cases involving determined medical conditions. Sample size restricted comparisons of plaintiff credibility to only some of the experts, including general practitioners, rheumatologists, physiatrists, orthopedic surgeons, and psychiatrists/psychologists testifying for the plaintiff, and for rheumatologists and orthopedic surgeons testifying for the defense.

The binary credibility rating system allowed for an estimate of the percentage of plaintiffs and experts to be rated as “credible”. According to this converted system, the percentage of plaintiffs considered “credible” are similar to the percentages of experts considered “credible”. A comparison of defense orthopedic surgeons’ and plaintiffs’ binary credibility ratings indicated a similarity in the percentage of plaintiffs and experts considered “credible”. These results do not support the fifth hypothesis, which predicted that expert testimony would be considered more credible than plaintiff testimony. However, the binary rating system converts individuals that were initially perceived by the judge to be only “partially credible” to be rated as “credible”. This converted rating system can result in a loss of information associated with perceived credibility. For example, the majority of plaintiffs (63%) were originally perceived by judges to be “partially credible”. According to the binary system, those same plaintiffs would be rated as “credible”, along with 28% of the plaintiffs initially rated as “credible”. Therefore, the converted binary system may be less representative of how judges perceived credibility, than the initial credibility rating scheme.
The sixth hypothesis stated that expert testimony would be considered equally credible by the court, regardless of the expert's familiarity with the plaintiff. Treating and forensic expert credibility were examined and revealed that although mean credibility ratings appeared higher for treating experts compared to forensic experts, the differences were not statistically significantly different. These results suggest that expert credibility was perceived to be similar by the judge, regardless of whether the experts were testifying for the plaintiff or the defense. However, when expert categories were combined across treating and forensic categories, treating clinicians were perceived to be significantly more credible than forensic experts. A comparison of treating and forensic experts using the converted binary rating system indicated that treating and forensic experts were considered equally credible. The individual expert category and binary credibility comparisons provide support for the sixth hypothesis. However, a comparison of the combined experts across treating and forensic categories does not support the sixth hypothesis.

**Limitations and Considerations for Future Research**

One limitation in this study is that because the case judgments involved a number of different judges, each judge differs in the amount and type of details reported in the case. This variation in details results in a lack of consistent information across case judgments. For example, one judge may describe the professional background of each medical expert, while another judge may not include that information. However, the large sample size of case judgments may have reduced the effects of this limitation. The fact that this is a retrospective study limits the ability to clarify information related to the variables, or gather more detailed information. The generalizability of these results is limited to claims pertaining to FM, but there may be some similarities with claims involving other undetermined chronic pain or medical conditions where there is no objective evidence of pathology. Time frame may also limit the generalizability of these results because of various factors, including changes to the provincial legislation of insurance systems, and changes in medical diagnostic information and practices.

A consideration for future research would be to follow the current case judgment sample through possible appeals processes that may have occurred since the initial trial. A follow-up of this sample would provide a clearer picture of the final outcome of each case in the litigation process.
In an attempt to further understand the factors associated with credibility in the courts, researchers could devise a scheme to code the reasons judges provided in their perceptions of witness credibility. Of course, a greater understanding of the factors related to credibility in the courts could be obtained in a prospective survey of judges, where they could provide details regarding their perceptions in cases involving claims for FM and other chronic pain conditions. Another direction for future research would be to examine whether there is an association between the gender of judges, plaintiffs, and variables such as credibility and award granted. A final consideration for future research would be to conduct a prospective study examining judges’ personal experience with pain and injury, and the possible role those experiences may play on their perceptions and decisions in similar cases.

A multidisciplinary approach has been useful in allowing for a broader focus, and resulting in a greater understanding of the results, than if only one perspective had been examined in this study. For example, an examination of plaintiff, surveillance, and expert credibility has demonstrated how credibility from various perspectives can affect the outcome of a claim. Therefore, this approach should be considered for future research.

In final summary, judges awarded disability claims for FM in the vast majority of cases (92%), and awards were not granted in only 7% of cases. Therefore, regardless of the ongoing debates in the medical community on the existence or diagnostic utility of FM, or whether FM is a compensable condition, the Canadian courts have tended to award FM cases with disability compensation. Plaintiff credibility was clearly paramount in this process, given the significant relationship between the judge’s perception of plaintiff credibility and amount of awards granted. It remains unclear as to what specific factors were involved in judges’ determination of credibility. The fact that the majority of plaintiffs (63%) were perceived by judges to be “partially credible”, suggests that determining credibility was not an easy task for judges. Judges were able to make a clear determination of plaintiff credibility in less than half of the cases. Credibility is crucial to the plaintiff in litigation and disability claims situations, but health care practitioners also require credible, reliable information in order to assist their patients (Craig & Badali, 2004), as do insurers in claims determinations. It would, therefore, be useful to examine ways of assisting the claimant, the medical practitioners, the insurance industry, and the courts in the determination of credibility and the accurate presentation of a condition such as FM.
Study 2: FM and Exaggerated Disability

Issues relating to fraud, malingering, and the deliberate misrepresentation of compensation claims were discussed in the insurance and legal sections of the introduction. These issues are of considerable importance from several perspectives, and have lead researchers to explore various methods aimed at the detection of exaggerated or malingered disability. In addition, the concepts of malingering and exaggeration have sometimes been applied interchangeably, even though they are separate concepts. Malingering is the intentional production of false or greatly exaggerated symptoms for the purpose of attaining some identifiable external reward (American Psychiatric Association, 1994). Exaggeration of symptoms, or deliberately poor effort during testing, describes the underlying behaviour without inferring motivation (Iverson, 2003).

In terms of base rates, Fishbain et al. (1999) conducted a detailed review of the literature examining malingering and disease simulation among the chronic pain patient population, and reported that malingering was present in 1.25 – 10.4%. Base rates of probable malingering or symptom exaggeration in a more recent study have been reported to be 29% for personal injury, 30% for disability, 35% for FM/chronic fatigue syndrome, 31% for chronic pain, and 19% for criminal involvement (Mittenberg et al., 2002). These base rates were obtained from estimates reported from a survey of the American Board of Clinical Neuropsychology members, so the accuracy of these estimates have not been confirmed. Unfortunately, this study did not differentiate the concepts of malingering from symptom exaggeration through much of the review; it is therefore unclear whether these concepts have been correctly identified with the associated base rates. The authors also did not differentiate the base rates for FM from chronic fatigue syndrome. Higher rates of probable malingering were reported for cases referred by defense counsel and insurers in civil matters, and in criminal cases referred by prosecutors. The authors concluded that these base rates and diagnostic impressions might be partially influenced by the selection of cases by the referral source.

Issues related to exaggeration, malingering, and the deliberate misrepresentation of disability level adds to the complexity of the claims process, especially for controversial disorders like FM. It is therefore essential to explore methods that could assist in determining the credibility of those diagnosed with FM. The focus of the second study is the examination
of techniques that could assist with the identification of exaggerated disability in patients with FM. To appreciate the factors associated with such techniques, an understanding of the components related to the identification of exaggeration will be presented in the following section. First, the perspective of the patient and the experience of chronic illness and chronic pain will be presented, followed by factors related to the compensation process. The literature on effort testing and malingering in relation to chronic pain and disability claims will also be presented, as well as methodological issues related to effort testing.

**Patient Perspectives**

After reviewing the literature, Ware (1999) reported that the experience of chronic illness and chronic pain has indicated that sufferers face delegitimation, or the systematic failure to confirm one’s illness perceptions by friends, family, and medical professionals. Bury (1991) noted that legitimization of a medical condition can be more problematic where the same symptoms occur in milder forms in the normal population. For example, FM symptoms such as body aches, sleep disturbance, and fatigue may be trivialized. In studies of FM and chronic fatigue syndrome, the invisible nature of the condition, lack of ‘objective’ evidence, and familiarity of the symptoms have all been found to contribute to the perceived lack of credibility of sufferers (Asbring & Narvanen, 2002; Henriksson, 1995; Soderberg, Lundman, & Norberg, 1999; Ware, 1992). It has been argued that people with chronic, invisible, and contested conditions are caught in a double-bind, in that they may communicate their pain in an attempt to have the condition taken seriously, but the lack of accepted cultural ways of doing so (Hilbert, 1984) may contribute to the delegitimation of their condition (Richardson, 2005). For example, the pain sufferer could be accused of whining or being psychologically unstable. In an attempt to understand the factors involved in establishing a positive identity, Richardson (2005) conducted individual interviews with female chronic pain patients. She found that these women strove to present themselves as hard working, in control, and sane; thereby refuting assertions of malingering, psychological instability, and complaining.

Back pain is a condition similar to FM, in that the cause is often unclear, and there is typically no visible damage or connection to any serious disease. Glenton (2002) examined the illness experiences among back pain sufferers and found they feared that those around them questioned the reality of their pain and suspected them of being malingers,
hypocondriacs, or mentally ill. In a subsequent study, Glenton (2003) examined issues related to back pain patients and the sick role. The sick role concept is based on an “acute” medical model of sickness (Crossley, 1998), where sickness is characteristically time-constricted, responsive to treatment, and related to physical rather than mental health. For long-term or chronic illnesses, the doctor-patient relationship is considered less central, and the role of patient responsibility is emphasized. Similar to the previous study, Glenton (2003) reported that back pain sufferers were fearful that doctors and others would distrust their motives and question the reality of their pain if no objective signs of their pain were found. Participants also expressed fear that a lack of treatment alternatives could be seen as a sign that they were not attempting to return to work. Participants also indicated that identification of objective physical signs of disease, positive diagnostic tests, and diagnoses are often welcome and seen as proof of their pain and suffering. A diagnosis also provides medical absolution from responsibility for the pain, and from accusations of malingering, hypochondria, or mental illness that may follow when proof is lacking. Diagnoses were seen as leading to explanations and suggested treatments, and central to the legitimization of sick benefits, such as sick leave, medication, and access to medical and disability benefits. Access to healthcare and treatment was also seen as a sign that recovery is desired, and the medical community accepts the suffering as valid.

One of Glenton’s (2003) main conclusions was that it is not the physical characteristics of back pain that represent the greatest threat to back pain sufferers, but the suspicion that the pain does not really exist. This conclusion was supported by findings from an earlier study examining the experiences of chronic pain patients, who indicated that central to their suffering was the desire to be believed and to have their pain legitimized by healthcare practitioners (Reid, Ewan, & Lowy, 1991). Similarly, others have noted the frustration associated with not being able to substantiate their pain and suffering by some form of visible proof (Walker, Holloway, & Sofaer, 1999), and the role that diagnostic testing has in the legitimization of pain and illness (Rhodes, McPhilips-Tangum, Markham, & Klenk, 1999). White et al. (2002) found that a newly diagnosed group of patients with FM did not worsen over time, and they concluded that the FM label may have allowed for more appropriate medical management. Finally, Glenton (2003) concluded that the inability to
offer chronic back pain sufferers a clear diagnosis or explanation does not appear to liberate the patient from the sick role, but rather prolongs dependence on the doctor.

Barsky and Borus (1999) provide a contrary view, similar to the views presented earlier by Hadler (1996), Aronoff (1999), and others. They refer to medically unexplained somatic symptoms and conditions (including FM) as a ‘functional somatic syndrome’ that are characterized more by symptoms, suffering, and disability than by disease-specific, demonstrable abnormalities. They suggest that beliefs, expectations, the sick role, and psychological distress become important in amplifying, maintaining, and perpetuating these symptoms and related disability levels. They conclude that the influence of psychosocial factors must be taken into account in understanding and treating these conditions. This view has been supported by reports of an association between the persistence of chronic widespread pain and features of somatization (e.g., McBeth, Macfarlane, Hunt, & Silman, 2001). Yet others argue that this position trivializes and dismisses complaints of the physically ill, thereby adding to the patient’s suffering (English, 2000; Hendrick, 2000).

The experience of chronic illness and chronic pain can leave the patient feeling invalidated and marginalized. Therefore, credibility and validation of the pain and suffering from healthcare providers and others signifies acceptance, and may lead to improved management of the condition. However, psychosocial factors such as beliefs and psychological distress level may have an impact on disability level. As well, access to disability compensation systems and other external factors associated with the condition may also influence disability level and management. Some of those influences are discussed in the next section.

**Compensation & Prevalence Rates**

White and Thompson (2003) examined the prevalence rate of FM in an Amish community, a population that is culturally distinct and isolated from the rest of North American society. This community also limits their exposure to media influences, and does not access provincial or federal disability compensation systems. Hence, if litigation plays a major role in the prevalence of FM, then FM prevalence rates should be lower in Amish populations. However, FM prevalence was higher in this population (7%) than in any other previously reported population in North America (approximately 1-5%). White and Thompson concluded that their findings provide support for the relative unimportance of
Credibility and Fibromyalgia

litigation and compensation on FM prevalence rates. These results are also supported by studies suggesting that FM prevalence rates are higher in countries like Pakistan and South Africa (Farooqi & Gibson, 1998; Lydell & Meyers, 1992), where compensation might be expected to be less available, compared to countries like Sweden and Denmark (Jacobsson, Lindgarde, & Manthorpe, 1989; Prescott et al., 1993), where compensation might be expected to be more available. Wolfe (2003) suggested that the FM examinations conducted in the Amish study may have been unreliable, given that the prevalence rate was higher than other North American rates. He further comments that the psychosocial symptoms of FM, which were not assessed in the Amish study, are more centrally important to the syndrome than the physical criteria of tender points.

Compensation has been associated with factors such as physical and emotional trauma. For example, Aaron et al. (1997) found that patients with FM whose condition was preceded by a physical trauma were significantly more likely to receive disability compensation than those without a trauma history. FM preceded by an emotional trauma was associated with increased healthcare utilization, and reports of greater functional disability, than those whose FM was preceded by a physical trauma. There were no differences in reported pain threshold levels or pain intensity as a function of trauma history between FM groups. Greenfield et al. (1992) reported similar findings for an association between FM that was preceded by a physical trauma and an increased likelihood of receiving disability compensation.

Some patients become involved in the disability claims process in an attempt to access compensation and treatment. Factors such as trauma history appears to influence disability level and healthcare utilization in patients with FM. Involvement in the claims process has also been associated with a variety of other factors, such as psychological and physical distress, as well as disability level. Studies related to these factors are summarized below.

Compensation Status & Chronic Pain

Teasell (2001) has suggested that stress from an adversarial compensation claim process may serve to exacerbate pain. Winning the claim may also become the patient’s focus and serve as a vindication of the injury. The compensation claim process may affect the
individual involved in various ways. This section will present those effects from a variety of viewpoints.

After reviewing the literature, Mendelson (1984a, 1992) reported that between 35-85% patients did not fully recover after settlement of their claims, and continued to receive medical treatment and remain disabled up to three years post litigation. Subsequent to his review, Mendelson (1995) conducted a followed up of 760 personal injury litigants and found that 75% had not returned to work after approximately 23 months following finalization of their cases. However, a group of patients could not be traced at follow-up, which brings the potential number of unemployed patients to 55%. Mendelson also reported that age played a role in return to work status, so that younger patients (Mean = 37 yrs) were significantly more likely to return to work prior to the finalization of claims than older patients (Mean = 42 yrs).

A meta-analysis of disability compensation and chronic pain studies revealed that receiving financial compensation was associated with a greater experience of pain and reduced treatment efficacy (Rohling, Binder, & Langhinrichsen-Rohling, 1995). The authors indicated that the reviewed study samples were representative of chronic pain patients seen in pain clinics and by psychologists, and suggested that patients who experience greater pain and greater difficulty in treatment may be more likely to seek and obtain compensation. Involvement in the compensation process has also been associated with greater emotional distress in several studies (Guest & Drummond, 1992; Hee et al., 2001; Turk & Okifuji, 1996). In a more recent review, Teasell (2001) found that involvement in the compensation process contributes to a lengthier recovery period and negatively influences response to treatment. He reported that there was no definitive evidence to suggest that compensation affects the incidence of injury or the development of chronic pain. Teasell further indicated that there were fewer and shorter claims when replacement wages were too low to live on, for either low or high-wage earners. The finding that claim rates and durations are influenced by the amount of compensation suggests that financial incentives play a role. Hall, McIntosh, Wilson, and Melles (1998) reported that, when people do not have insurance, leisure time activities were regarded by most to be responsible for patients’ pain, while 90% of those insured claimed the pain to be work-related.
Some authors have reported limited or no differences on pain severity ratings or psychological diagnoses in relation to compensation status (Leavitt, Garron, McNeill, & Whisler, 1982; Melzack, Katz, & Jeans, 1985; Mendelson, 1984b). However, Leavitt and colleagues did find that psychologically healthy patients on compensation who had objective evidence of organic disease reported increased pain sensations. One study reported that although involvement in the compensation process was associated with greater life disruptions due to pain, compensation status did not affect pain severity or psychological distress (Tait, Margolis, Krause, & Liebowitz, 1988).

Whether the compensation process is associated with negative influences or not, the presence of financial incentives and the potential for malingering remain clear. The ability to detect exaggeration and malingering is an important consideration from an economic and patient care perspective.

**Symptom Exaggeration & Chronic Pain**

As summarized earlier, involvement in the compensation process may affect an individual in various ways, but it is clear that the process can be very distressing. Hadler (1998) points out that people with FM-related symptoms may be involved in tort or disability determinations, and struggle with their ability to rehabilitate, while attempting to prove their illness in order to access treatment and compensation. Hence, the question for some clinicians is whether patients with FM provide an accurate presentation of their condition, or whether it is influenced by claims issues and access to treatment.

Waddell and colleagues described a group of eight physical signs in low back pain patients that were frequently found in chronic pain patients (Waddell, McCulloch, Kummel, & Venner, 1980; Waddell, Pilowsky, & Bond, 1989). These signs were identified as nonorganic, and thought to be predictors of psychological problems and distress. Prior to Waddell’s findings, Frost (1972) described these signs as representing hysteria or malingering. Waddell has suggested that these signs may be an indicator of abnormal illness or pain behaviour (Waddell et al., 1989, Waddell, 1987). As a result, there has been considerable controversy over the meaning of these signs. In a recent review, Fishbain and colleagues concluded that there was very little association between malingering, secondary gain, and the presence of these signs (Fishbain, Culter, Rosomoff, Rosomoff, 2004).
In terms of psychological and personality assessment measures, the Minnesota Multiphasic Personality Inventory – 2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) has been used to examine chronic pain populations. Slesinger, Archer, and Duane (2002) found higher levels of depressive symptomatology, anxiety, and somatic complaints in comparison to a normal control group. These findings are consistent with those reported by Love and Peck (1987). The MMPI – 2 has also been examined in the context of detecting exaggeration and malingering in chronic pain patients. In particular, the “Fake Bad Scale” (FBS) was developed to assess simulation or exaggeration of emotional distress injuries or disability in personal injury litigants (Lees-Haley, English, & Glenn, 1991). This scale has been reported to detect the exaggeration of psychological problems in litigants presenting with emotional trauma and pain (e.g., Greiffenstein, Baker, Axelrod, Peck, & Gervais, 2004), and the exaggeration of physical and psychological problems (Larrabee, 1997, 1998). A study examining the FBS in a sample of personal injury litigants concluded that high FBS scores could be affected by exaggeration and litigation stress (Tsushima & Tsushima, 2001).

Meyers, Millis, and Volkert (2002) examined several validity scales on the MMPI-2, including the FBS, in a sample of chronic pain patients. They found that the FBS correctly identified all of the nonlitigant patients, but only 15% of the patients in litigation. They combined the validity scales into one total weighted score, and the weighted score increased the identification rate of patients in litigation to 36%. A significantly different profile was found on the validity scales of chronic pain patients in litigation, compared to nonlitigants. The authors concluded that a single validity scale may produce more false positives than the weighted score of multiple validity scales.

One study examining the psychometric characteristics of the FBS reported that patients with a broad range of physical problems, such as chronic pain, tend to produce high scores on this scale simply by reporting their physical problems (Butcher, Arbisi, Atlis, & McNulty, 2003). They also reported a bias toward classifying women as malingeringers with the FBS. Therefore, the authors suggested that the FBS should not be used to detect malingering of chronic pain because of unacceptably high false positive rates. This is not surprising, given that the FBS was not designed to detect exaggeration in chronic pain patients. The authors
also noted that many of the published studies on the FBS do not have a clearly determined “malingered” and “nonmalingered” sample on which to verify the classification success.

In a review of the literature, Arbisi and Butcher (2004) suggest that the malingered versus nonmalingered dichotomy fails to acknowledge the changing nature of both the pain experience and the motivation to alter self-report. The clinician is required to assess the degree of self-reported pain, and account for the possibility of distortion in that process. Arbisi and Butcher suggest that the MMPI-2 can provide useful information in the degree to which the individual has altered their presentation, and the psychological context of the pain. A comparison of symptomatic status and level of psychological distress, relative to other patient groups, and elevations on some of the validity scales can provide information regarding accuracy of self-report, patient’s response attitudes, and level of cooperation with the assessment. They further indicate the need for a validity scale aimed at the malingering of somatic complaints and physical pain on the MMPI-2. Iverson and Lange (in press) point out that the Butcher et al. (2003) study has been criticized for methodological limitations, and they further indicate that the literature tends to support the use of the FBS in detecting exaggeration in personal injury cases.

Literature related to the MMPI-2 suggested that chronic pain patients tend to produce a distinct profile on some of the clinical scales, with elevations on scales measuring depression, anxiety, and somatic complaints. Although, there is considerable debate over the utility of the FBS in detecting malingering, there is support for the FBS in detecting exaggeration in personal injury cases. It is clear that, as suggested by Butcher et al. (2003), the FBS needs to be examined in non-experimental situations in order to determine it’s utility in detecting malingering.

Several of the validity scales included in the Personality Assessment Inventory (PAI; Morey, 1991) have also been examined in terms of determining the accuracy of symptom reporting and possible malingering of mental health problems. Those validity scales include the Negative Impression Management, Positive Impression Management, Malingering Index, and Rogers Discriminant Function. Two studies have reported that the PAI is limited in the detection of exaggerated and malingered depression, with a misclassification of approximately half of a sample of subjects feigning depression (Rogers, Ornduff, & Sewell, 1993; Rogers, Sewell, Morey, & Ustad, 1996). The misclassification rates were even higher
in the identification of feigned generalized anxiety. However, these studies found greater success for detecting feigned schizophrenia on the PAI. The literature also suggests that Post-Traumatic Stress Disorder can successfully be feigned on the PAI (e.g., Calhoun, Earnst, Tucker, Kirby, & Beckham, 2000; Liljequist, Kinder, & Schinka, 1998). Using the recommended cutoff scores on several of the validity scales, less than half of the subjects that were simulating the disorder were correctly classified. Several of the PAI validity scales used in combination were found to be moderately successful at differentiating individuals instructed to simulate non-specific mental illness from actual patient populations and healthy controls (e.g., Bagby, Nicholson, Bacchiochi, Ryder, & Bury, 2002; Blanchard, McGrath, Pogge, & Khadivi, 2003; Poythress, Edens, & Watkins, 2001). However, the PAI validity and clinical scales have not yet been explored in the detection of exaggerated pain and disability.

There have been attempts to determine the accuracy of symptom reporting for specific mental health problems in the absence of validity scales. For example, Lees-Haley (1989) explored the ability to feign depression on the Beck Depression Inventory and found that college students were very successful. The author cautioned against the use of such instruments in situations where deliberate exaggeration or manipulation of responses is a possibility. Steffan, Clopton, and Morgan (2003) constructed and evaluated the Malingered Depression scale (Md); a new validity scale for the MMPI-2 aimed at distinguishing malingered depression from genuine depression. They found that the Md scale had a hit rate (correctly identification) of 91% for genuinely depressed participants, and a negative predictive power (correctly classifying participants feigning depression) of 97%.

Leavitt (1985) explored pain reports in the context of malingering and exaggerated disability and found that simulators used words reflecting intense pain or affective distress more frequently than back pain patients. Initial results revealed a 66% accuracy rate in predicting pain simulators, and a 93% accuracy rate in predicting pain patients, while cross-validation results revealed a 59% accuracy rate for simulators, and an 83% accuracy rate for patient controls. One study examined the impact of response bias on three self-report pain assessment measures (Multidimensional Pain Inventory, Coping Strategies Questionnaire, and Pain Beliefs and Perceptions Inventory), and compared students simulating the response of someone either “coping well” or “coping poorly” (Robinson et al., 1997). They found that
subjects could convincingly portray themselves on each assessment instrument according to the instructional set, and that subject response patterns in the "coping poorly" condition were indistinguishable from a chronic pain patient group. They concluded that validity indicators are essential for detecting biased responding on these measures. Wallis and Bogduk (1996) found that individuals simulating physical disability scored significantly higher on the Symptom Checklist-90-Revised (Derogatis, 1983), the McGill Pain Questionnaire (Melzack, 1975), and a visual analogue pain scale, than a group of whiplash patients. The authors concluded that the pattern of over-reporting for individuals simulating whiplash was an indication that it is difficult to fake the psychological profile of a whiplash patient. Larrabee (2003) examined three self-report pain assessment measures (McGill Pain Questionnaire, Pain Disability Index, and Modified Somatic Perception Questionnaire) among a group of mild head injury litigants with pain complaints. One group of litigants with no objective findings were considered to be probable malingers, and another group were considered nonmalingers. The probable malingers scored significantly higher on all three of the pain scales, compared to a group of pain patients. The Modified Somatic Perception Questionnaire was found to be better than the remaining measures at classifying individuals in the probable malingerer group, with a sensitivity (correctly identified exaggerators) of 90%, and specificity (correctly identified non-exaggerators) of 90%. None of the pain measures mentioned above included validity scales.

McGuire and Shores (2001) investigated the effectiveness of a pain assessment measure that included a validity measure for detecting exaggerated pain complaints (Pain Patient Profile; Tollison & Langley, 1995). They found that pain simulators could be differentiated from pain patients by excessive elevations on the clinical scales, but not on the validity scale. A related study using the same measure found that a group of pain patients instructed to exaggerate their pain, and a general clinical group instructed to simulate malingered pain, could be differentiated from a control group of pain patients responding normally by their elevated clinical scale scores (McGuire, Harvey, & Shores, 2001). For example, there was a sensitivity of 85% and a specificity of 87% on the somatization scale, and a sensitivity of 77% and a specificity of 77% on the depression scale, between the pain group exaggerating their pain and the control group, whereas, the validity scale had a sensitivity of 69% and a specificity of 81% between these two groups.
In situations involving compensation or secondary gain, patients with secondary gain factors do not consistently respond differently than patients with no secondary gain factors. For example, Fishbain and colleagues compared two groups of chronic pain patients, one group with secondary gain factors and another group with no secondary gain factors, pre and post treatment (Fishbain Cutler, Rosomoff, & Steele-Rosomoff, 2002). The secondary gain group did not differentially respond on an illness behavior questionnaire and scale designed to detect conscious illness exaggeration (Conscious Exaggeration Scale; Clayer, Bookless, & Ross, 1984) compared to the control group. The authors concluded that the illness exaggeration scale was either not a valid instrument for identifying exaggeration in patients with chronic pain, or that patients were not exaggerating. Similar results were reported in an earlier study using the same exaggeration scale, comparing a compensation and non-compensation group of chronic non-cancer pain patients (Mendelson, 1987).

Meyers and Diep (2000) investigated the ability of six neuropsychological tests to detect exaggerated cognitive impairment in a sample of litigant and nonlitigant chronic pain patients reporting decreased mental functioning. They reported that 29% of the litigant pain patients failed two or more of the six neuropsychological tests, compared to none of the non-litigant pain patients. They concluded that the measures they included could be used to assess exaggeration of cognitive symptoms in chronic pain patients.

Fishbain et al. (1999) conducted a detailed review of the research examining malingering and disease simulation techniques among a chronic pain patient population. Specifically, they found that studies of facial expressions, sensory testing, questionnaires, isometric testing (e.g., hand grip), and clinical examination could not reliably detect malingering within a chronic pain patient population. They further indicated that isokinetic strength testing showed potential for the discrimination between maximal and submaximal effort, and between best and malingered effort. Isokinetic testing has limited value in the detection of malingering for many medical conditions, such as FM, because physical strength is not a defining feature of the condition. Fishbain and colleagues concluded by stating that there are currently no reliable methods to identify malingering in patients with chronic pain. A recent review of muscle testing techniques for the determination of sincere effort also indicates that although some techniques are promising, they have not been sufficiently validated for clinical application (Robinson & Dannecker, 2004). The authors further caution
that sources of variability in testing (e.g., pain and fear) have been previously neglected, and are crucial to the clinical application of these methods.

Fishbain et al. (1999) suggest that the frequencies of actual malingering cases within different populations are almost nonexistent, and the available data are not consistent. Many studies include litigant patient populations as their "malingering" or "exaggeration" group. Of course, there is no evidence that all litigants exaggerate or malinger. Iverson and Lange (in press) suggest that very little is known about a litigant's underlying motivation to engage in deliberate exaggeration. Bianchini, Greve, and Glynn (2005) indicate that accurate estimates of sensitivity and specificity require accurate group assignment, and that "one must be very confident that only persons who are malingering have been included in the malingering group." (p. 406). Although multiple criteria have been suggested as an essential part in increasing the accuracy of assignment to a malingerer group (Slick, Sherman, & Iverson, 1999), those criteria have not been consistently applied in the malingering research. Also, Iverson (unpublished manuscript) has suggested that a conceptual confusion exists amongst researchers regarding the similarities and differences between the constructs of exaggeration and poor effort. The distinction between them has often been blurred, which can be misleading in the malingering research. Therefore, malingering research involving estimates of the prevalence of malingering should be examined carefully.

**Effort Testing**

Procedures for detecting poor effort and the use of effort testing have been an essential component in neuropsychological assessments for many years. Iverson and Lange (in press) describe two main approaches that have been used to detect poor effort in neuropsychology. The first approach involves the examination of performance patterns on specific neuropsychological tests or test batteries. Performance is considered to be suspect if it is (a) uncommon in healthy people, (b) uncommon in patients with psychiatric conditions, brain injuries, or brain diseases, yet (c) relatively common in people known or thought to be exaggerating or malingering. A second approach has been the development of specialized tests designed to detect poor effort. In effort testing, the basic assumption is that if an individual has no memory for the test material, the lower limits of correct performance will be at chance levels (Iverson & Frazen, 1998). An individual is considered to be responding with poor or suboptimal effort when the number of errors falls significantly below that
expected by chance. Cutoff scores, based on performances of patients with well documented brain dysfunction who are not in litigation, also provide important comparison data (Iverson, 2003). Iverson and Lange (in press) suggest that although these tests actually measure poor or suboptimal effort; by inference, they are a measure of exaggerated cognitive impairment.

Iverson (2003) indicates that very little is known about an individual’s motivation to engage in deliberate exaggeration of symptoms or problems. He further points out that the presence of exaggeration or poor effort on testing should not be automatically equated with malingering. A number of differential diagnoses and alternative explanations should be considered in the evaluation process. Some differential diagnoses to consider are somatoform disorders, depression, factitious disorder, and hypochondriasis (Iverson & Binder, 2000).

Litigation can be very stressful, and the need for multiple evaluations is common. Iverson (2003) suggests that the plaintiff may encounter very different environments in the health care system than a person not in litigation. For example, some health care workers may have a vested interest in the assessment and rehabilitation of the plaintiff, while others may have a vested interest in doubting the veracity of the patient’s reported problems. In addition, family member, lawyers, and expert witnesses may also have competing interests (Iverson & Lange, in press). The result of these differing interests may result in hostility or sympathy from others, depending on their interests. The plaintiff may behave differently as a result, and may not be aware of the changes in their behaviour. Regardless of the motives behind exaggerated symptoms or poor effort, it is essential to determine whether the assessment of an individual’s condition is accurate.

The research examining effort tests strongly suggests that a substantial minority of people involved in personal injury litigation or disability claims exaggerate cognitive impairment. These results have been found with a number of patient groups, including people with mild head injuries (Binder, 1993; Grote et al., 2000), chronic fatigue syndrome (van der Werf et al., 2000), chronic pain following whiplash injuries (Schmand et al., 1998), and chronic pain with sensory loss (Greve, Bianchini, & Ameduri, 2003). Gervais et al. (2001) found similar results in a group of patients diagnosed with FM, whereby 35% of those involved in compensation or disability claims failed the effort tests, in contrast to only a 4% failure rate for those not involved in disability claims.
The effort testing research has clear implications relating to the credibility of patients' complaints and their compensation claims. Of course, determining the nature and extent of disability, and the veracity of symptom complaints, is necessary for proper adjudication within the tort law system, and proper insurance processing of disability claims. It is noteworthy to point out that, in general, effort testing research currently meets the Daubert criteria for acceptable expert evidence in certain clinical populations (Green, Lees-Haley, & Allen, 2002; Thompson, 2002; Tombaugh, 2002), and could be enhanced by further exploration within a FM population.

Methodological Considerations in Effort Testing

A great deal of research has been conducted on objective methods for evaluating the validity of test results, with a major emphasis on the use of effort tests, combined with an evaluation of the consistency between various types of evidence (e.g., Iverson & Binder, 2000). The approach to effort testing has evolved over the past decade in neuropsychology, and some of those advances relevant to the assessment of chronic pain patients are described below.

It has been suggested that reliance on a single effort test can produce too many false positive and false negative errors (Greve et al., 2003; Pankrantz, 1988). For example, Gevais, Rohling, Green, and Ford (2003) compared the performance on three effort tests and found that more than twice as many people failed the one involving verbal stimuli than the one involving visual stimuli, and that these tests varied substantially in their sensitivity to response bias. These results suggest that a more cautious approach in examining effort would be to include multiple measures of response bias. The use of multiple validity measures has been tested and proven successful by improved correct classification rates. Meyers and Diep (2000) used multiple neuropsychological measures and found that close to one third of chronic pain patients in the litigant group failed two or more validity checks included in these tests, compared to none of the patients in the nonlitigant group. Meyers et al. (2002) combined several validity scales on the MMPI-2 to assess malingering in chronic pain patients. They were able to correctly classify 100% of the nonlitigant patient group, and 86% of the litigant patient group. Meyers and Volbrecht (2003) also found that several validity checks on multiple neuropsychological tests together resulted in an increased ability to
Credibility and Fibromyalgia

correctly identify litigant and nonlitigant groups (including a chronic pain patients). Failure on the validity checks on any two of these tests was suggestive of poor effort in this study.

In terms of comparison groups, most malingering studies have utilized a design where participants simulate the disorder under investigation. It has been recommended that a clinical comparison group for the disorder under investigation be included in simulation studies, as opposed to ‘normal’ controls (Rogers, 1997). Resnick (1995) argues that simulating participants in previous studies have usually not been from clinical populations, and he refers to these individuals as ‘pure malingerers’. He further indicates that most malingerers encountered in clinical practice do have some level of impairment, but exaggerate its severity, and are actually ‘partial malingerers’. Most simulation studies have not taken this distinction into account (McGuire et al., 2001).

In accordance with the advances in effort testing described above, multiple measures of effort have been employed in Study 2, including three effort tests, as well as validity scales on a personality assessment measure. Study 2 involves a within-subjects design, so that the comparison group also represents the medical condition under investigation, patients with FM.
Hypotheses

The purpose of this study was to examine the phenomenology of exaggeration and the deliberate portrayal of excessive disability in patients with FM. This study has both descriptive and experimental components. A multimodal, comprehensive assessment was completed twice with each participant within two conditions; namely, the “normal” (full effort) condition and the “exaggeration” (excessive disability) condition. Patients who currently met the ACR diagnostic criteria for FM were asked to complete all tests and measures in two conditions, in a counter-balanced order. In one condition, they were asked to give their normal effort, and in another condition they were asked to exaggerate their disability in a manner that would indicate high levels of pain and disability. This is a repeated measures, counterbalanced analog experimental design in which patients served as their own controls. Two specific hypotheses are listed below.

1) Patients will exaggerate their disability in multiple domains (e.g., psychological, cognitive, physical, and daily functioning). Specifically, patients will perform significantly worse on all measures when in the exaggeration condition.

2) Using combinations of test scores, a majority of patients instructed to exaggerate will be identified correctly.
Method

Participants

Fifty-seven patients from two private rheumatology practices and one outpatient hospital program participated in the study. Those patients previously diagnosed with FM were sent a letter from the clinics informing them of the project. Patients attending the outpatient hospital program received a letter from the student investigator informing them of the project. The letter described to them that the nature of the project was to examine issues related to FM and soft tissue pain, and would involve the completion of a number of tests and questionnaires. The letter also described the time requirements involved, and offered an honorarium of $60 for participation. Following the letter, prospective participants were contacted by phone to confirm their willingness to participate and appointment times were scheduled.

Participants were excluded from the study if they did not meet diagnostic criteria (described below), or if they had been previously diagnosed with a significant co-existing rheumatic disease. Participants were also excluded if they were not able to travel to the testing centre, were not fluent in English, or did not have corrected-to-normal vision. Two individuals were excluded because they did not meet diagnostic criteria, and one individual was excluded because she was not fluent in English.

The vast majority of the sample was women (94%). Their average age was 51.4 years (SD = 12.8, Range = 17-75), and their average education was 13.5 years (SD = 2.4, Range = 7-20). The breakdown of marital status was as follows: 54% married, 5% commonlaw, 4% separated, 15% divorced, 5% widowed, and 17% single. The breakdown of the sample by ethnicity was as follows: 81% Caucasian, 7% Indian, 6% Asian, 4% European, and 2% other. Their employment status at the time of the study was as follows: 26% full-time employment, 17% part-time employment, 7% unemployed, 35% disabled, 15% retired, 1% student, and 1% other.

Diagnostic Assessment

All participants were diagnosed with FM by a rheumatologist on the basis of the ACR criteria (Wolfe et al., 1990). Those criteria include: (a) the presence of widespread pain in the four body quadrants lasting three or more months, and (b) pain and tenderness reports for at least 11 of the 18 recognized tender points upon digital palpation. A general medical
screening examination was also performed, in order to determine any co-existing medical conditions. Participants completed an information sheet related to demographic and medication information prior to the diagnostic assessment. The medication information was reviewed by the rheumatologist to ensure that the participant was on no medications that might interfere with an accurate evaluation of tender points or indicate a diagnosis that would exclude FM (e.g., taking methotrexate would suggest the presence of rheumatoid arthritis). Those who met the ACR criteria for FM and did not possess any significant co-existing medical conditions or exclusion criteria listed above were included in the study. On average, the participants had 13.8 tender points (SD = 2.12, Range = 11-18) and less than one control point.

Measures

A variety of measures were incorporated in the experimental sessions to examine response patterns from medical, functional, psychological, and symptom validity perspectives. Those measures are listed below.

Beck Depression Inventory – Second Edition (BDI-II) – This is a 21-item self-report test that measures the severity of key symptoms associated with clinical depression (Beck, Steer, & Brown, 1996). Participants are required to endorse one of four options for each item. Scores from 0 to 3 are applied to each option, with higher scores indicating more severe symptoms. Individual item scores are summed for a maximum total score of 63.

The BDI-II is an updated version of the BDI-1A (Beck, Rush, Shaw, & Emery, 1979). It was altered to correspond to criteria from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 1994) for major depressive disorder and to improve the content validity of the instrument. The previous version was believed to overestimate the severity of depression in medical patients, as a result of the somatic items inflating the total score by somatic complaints unrelated to mood (e.g., Morley, Williams, & Black, 2002; Wesley, Gatchel, Garofalo, & Polatin, 1999). Hassett and colleagues have suggested that the modifications of the BDI-II make it less sensitive to medical factors, resulting in a more appropriate measure of depression in patients with chronic pain (Hassett et al., 2000). The BDI-II has been used with a number of patient populations, including patients with minor medical conditions, FM, and chronic pain (Arnau, Meagher, Norris, & Bramson, 2001; Hassett et al., 2000; Viljoen, Iverson, Griffiths,
The BDI-II manual reports high internal consistency and test-retest reliability. In terms of validity, moderate to high convergent validity with other measures of depression, and high diagnostic discrimination with various outpatient samples are also reported in the manual.

Multidimensional Pain Inventory – Version 1 (MPI-1) – The MPI-1 is a 52-item self-report questionnaire divided into three parts (Kerns, Turk, & Rudy, 1985; Rudy, 1989). Each part contains several scales. The first part was designed to evaluate five dimensions of the pain experience, including patients' reports of pain severity, perceptions of how pain interferes with various areas of life, appraisal of the amount of support and concern from significant others, perceived self-control, and negative mood and distress. The scales measuring those five dimensions are pain severity, interference, support, life control, and affective distress. The second part was designed to assess patients' reports of the frequency of a range of responses by significant others to their displays of pain. The three scales measuring those reports are negative, solicitous, and distracting responses. The third part assesses patients' reports of their participation in common daily activities, including scales for household chores, outdoor work, activities away from home, and social activities. The final scale is General Activity Level, and a score for this scale is calculated based on scale scores from the third part of the test. Each question is scored on a 7-point scale (0-6), with 6 representing the most severe response.

Kerns et al. (1985) reported satisfactory internal consistency (Cronbach's range from 0.70 – 0.90) and good test-retest stability of the 13 scale scores over a 2-week time period (0.62 - 0.91). Similar test-retest reliability has been reported by others over the same time period (Bergstrom et al., 1998; Lousberg, van Breukelen, Schmidt, Arntz, & Winter, 1999). The authors also reported good convergent and discriminant validity of the 12 specific scales. The MPI-1 has been used in both chronic pain and FM patient samples (Bernstein, Jaremko, & Hinkley, 1995; Lousberg, Groenman, & Schmidt, 1996; Turk, Okifuji, Sinclair, & Starz, 1996).

The MPI-1 is based on patterns of responding, and patients can be classified into one of three adaptation or coping profile categories (dysfunctional, interpersonally distressed, adaptive coper; Turk & Rudy, 1990; 1992). A recent study examining the stability of this classification system reported poor test-retest reliability with female FM patients (Broderick,
Junghaenel, & Turk, 2004). Approximately one-third of the FM patients changed coping profile categories over a 2 – 4 week time period. However, test-retest stability of the individual scale scores of the MPI were within an acceptable range (0.70) for these FM patients. The authors concluded that the MPI classification categories may not be stable, trait-like patterns for many chronic pain patients. Hence, the classification system was not used in the current study.

Fibromyalgia Impact Questionnaire (FIQ) – The FIQ is a 19-item self-report inventory that is used to assess the current health status of patients with FM (Burckhardt, Clark, & Bennett, 1991). The first 10 items are evaluated collectively to provide a single score of physical disability, which focuses on the ability of patients to perform large muscle tasks. The next two items assess work status, and are not included in the final score. The remaining items assess impact of illness on a variety of areas (job, pain, fatigue, sleep, stiffness, depression, and anxiety), and are measured on visual analog scales. The total FIQ score was calculated according to the method proposed by Burckhardt et al. (1991). The range of scores is 0-80, with higher scores indicating greater impact of FM on the patient’s life. Physical functioning related items were reported to be highly reliable (0.95) on a six week test-retest interval, whereas pain related items were in lower (0.56) but acceptable test-retest (Pearson’s r) correlations (Burckhardt et al., 1991). The authors suggested that the moderate correlations for pain related items may be a reflection of the symptom fluctuations and unpredictable nature of FM. In terms of construct validity, significant correlations were reported between the FIQ and a comparable measure, indicating convergent validity of both instruments (Burckhardt et al., 1991; White & Harth 1996).

Pittsburgh Sleep Quality Index (PSQI) – The PSQI consists of 19 self-rated questions that assess sleep quality and quantity (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). The questions are grouped into seven component scores: subjective sleep quality, sleep latency, duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. The seven component scores are then summed to yield a global score (maximum score = 21). Higher scores indicate worse sleep quality.

The authors reported that the component scores as well as individual questions of the PSQI had a high degree of internal consistency, indicated by a reliability coefficient (Cronbach’s α) of 0.83. The global and individual scores were considered to be stable across
time. The test-retest interval was an average of 28 days, and the reliability coefficient for the global score was 0.85. The exception to this outcome was medication use in control subjects, which showed no correlation between test times. A global PSQI score greater than 5 has been reported to discriminate between good and poor sleepers, with a sensitivity of 89.6% and specificity of 86.5% in that discrimination (Buysse et al., 1989; Buysse et al., 1991). Similarly good test-retest reliability and construct validity have been reported by others (Backhaus, Junghanns, Broocks, Riemann, & Hohagen, 2002; Beck, Schwartz, Towsley, Dudley, & Barsevick, 2004). The PSQI has also been used in a variety of subject samples, including patients with various somatic disorders (Carpenter & Andrykowski, 1998), psychiatric disorders (Buysse et al, 1989), and cancer patients (Beck et al., 2004).

Oswestry Disability Index – version 2.0 (ODI-2.0) – The ODI-2.0 is a 10-item, self-report instrument assessing the level of pain interference with physical activities (Baker, Pynsent, & Fairbank, 1989). It was designed to assess a broad domain of functioning, including personal care, lifting, walking, sitting, standing, sleeping, sexual activity, social life, and pain intensity. Individuals are asked to endorse a set of statements that describes their current level of functioning or limitations. Each question is scored on a 6-point scale (0-5), with 0 representing no limitation and 5 representing maximal limitation. The questions are summed for a total score of 50, which is converted into a percentage.

The ODI-1.0 version was designed specifically for low back pain patients (Fairbank, Couper, Davies, & O’Brien, 1980), while the 2.0 version was revised and recommended for assessment of more general pain conditions (Baker et al., 1989). Researchers have reported an acceptable degree of internal consistency for version 2.0, with Cronbach’s \( \alpha \) ranging from 0.76 (Fisher & Johnson, 1997), to 0.87 (Kopec et al., 1996). Although test-retest reliability of version 2.0 was considered to be good (\( r = 0.89 \)), an increase in the time interval may lower test-retest reliability because of a natural fluctuation in symptoms (for a review see Fairbank & Pynsent, 2000).

Personality Assessment Inventory (PAI) – This is a 344-item self-report inventory designed to assess personality and psychopathology (Morey, 1991). The PAI items are answered on a 4-point, Likert scale (false, not at all true, slightly true, mainly true, and very true). There is a written and computer version of the PAI. The written version was used in the current study. The PAI provides scores on 11 clinical scales (Somatization, Anxiety,
Anxiety-Related disorders, Depression, Mania, Paranoia, Schizophrenia, Borderline Features, Antisocial Features, Alcohol Problems, and Drug Problems), and 31 related subscales. All clinical scales were developed to be compatible with DSM-IV clinical criteria, and each scale includes items that address the full range of severity of each clinical construct (Morey & Henry, 1994). In addition, there are four scales that assess response style and validity, and five treatment consideration scales (aggression, suicidal ideation, stress, nonsupport, and treatment rejection). The four validity scales (Inconsistency, Infrequency, Positive Impression Management (PIM), and Negative Impression Management (NIM) address the possible influence of biased or careless response sets. The author has added other indices for assessing response bias, which include the Defensiveness Index (DEF), the Malingering Index, and the Rogers Discriminant Function (RDF).

Internal consistency reliability has been reported to be good for both the full clinical scales and related subscales (Boone, 1998; Boyle & Lennon, 1994; Morey, 1991). However, the author reported lower internal consistency estimates for the Inconsistency and Infrequency scales, and indicated that because these scales measure response styles, they are expected to involve greater variability. The author also reported good test-retest reliability in the stability of clinical scale configurations over time (approximately one month). The clinical scales were designed to measure a wide range of severity of each construct, reported to have good internal consistency, as well as good convergent and discriminant validity (Morey, 1991). The reliability and validity studies include both inpatient and outpatient psychiatric samples, individuals from medical settings and substance-abuse programs, and various other samples (Morey, 1991; Morey & Henry, 1994). Research has supported the discriminative power of several of the validity scales and associated indices, including the RDF, PIM, and DEF (Bagby, Nicholson, Bacchiochi, Ryder, & Bury, 2002; Cashel, Rogers, Sewell, & Martini-Cannici, 1995; Morey, 1991; Morey & Lanier, 1998; Peebles & Moore, 1998; Rogers, Sewell, Morey, & Ustad, 1996). Some studies indicate marginal discriminative power for some scales, including the NIM, ICN, and INF (Rogers, Ornduff, & Sewell, 1993; Rogers et al., 1996).

Test of Memory Malingering (TOMM) – The TOMM is a visual recognition memory test that incorporates line drawings of common objects as stimuli (Tombaugh, 1996). The same items are presented in two learning trials, and each learning trial has two phases: a
study and a test phase. In the study phase, subjects are presented with a series of 50 stimuli. During the test phase, subjects are presented with the same 50 study stimuli as well as several “distractor” stimuli. They must identify the study stimuli. The written version of this test was used in the current study. This test was designed to detect exaggerated memory impairment. Scores are compared with cutoff scores that were validated in a series of experiments and through comparison with data obtained from patients with well-documented brain dysfunction (Rees, Tombaugh, Gansler, & Moczynski, 1998; Tombaugh, 1997). The TOMM has been reported to be a robust test that is relatively unaffected by age, education, a wide variety of neurological impairments (Tombaugh, 1996, 1997), depression, or anxiety (Ashendorf, Constantinou, & McCaffrey, 2004; Rees, Tombaugh, & Boulay, 2001). It has also been reported to have high sensitivity and specificity for distinguishing suspected malingerers and those simulating memory loss from normal controls (for a review see Tombaugh, 2002).

Victoria Symptom Validity Test (VSVT) – The VSVT is a 48-item computerized digit recognition, symptom validity test (Slick, Hopp, Strauss, & Thompson, 1997). Participants initially view a five-digit study number during the “Study Trial”, followed by a “Retention Interval”, during which time the computer screen is blank. Then the “Recognition Trial” is presented, during which time the initial five-digit number and a new five-digit number are presented simultaneously and the participant is instructed to choose which of the two five-digit numbers was presented during the Study Trial (forced choice). The task involves increasing levels of perceived difficulty associated with lengthening the Retention Interval from 5, to 10, to 15-second intervals. These intervals are presented in three blocks of 16 items each, for a total of 48 items.

This test was developed to identify poor effort during a neuropsychological evaluation. The basic premise is that if a patient has no memory of the stimuli, then the lower limits of correct performance will be at chance levels (i.e., approximately 50% correct). Empirically derived cutoff scores also have been developed. Several studies have reported the ability of the VSVT to detect poor effort (for a review see Thompson, 2002).

21 Item Test – This is a one-trial, orally presented word list containing 21 nouns (Iverson, 1998; Iverson, Franzen, & McCracken, 1991; 1994). Immediately following presentation of the list, participants are asked to freely recall the words. This is followed by a
two-alternative forced-choice task in which the participant is asked to select the target word from a target-distracter word pair. Four scores are derived from this test. The Free Recall score is the total number of words correctly recalled from the free recall portion of the test. The Forced Choice score is the total number of correct responses on the forced-choice portion of the test. The Inconsistency score is the number of words correctly recalled during the free recall portion that were not recollected on the forced-choice component. Finally, the Greatest Consecutive Misses score is a calculation of the greatest number of consecutive words not recalled on the forced-choice component. This test was designed to measure poor effort. Patient’s scores can be compared to chance performance levels and to empirically derived cutoff scores.

**Procedures**

After consent was obtained verbally, participants were scheduled for the medical screening assessment and first experimental session. Upon arrival, written consent was obtained from each participant, and limits of confidentiality were reviewed by a research assistant. Participants were informed that data from the study would remain confidential except under the following conditions: (a) if the information provided leads the researcher to believe the participant is a danger to themself or others, (b) if the information provided leads the researcher to believe a child is in danger, or (c) if a court subpoenas the study records. Once the consent form was signed, participants completed the information sheet related to demographic and medication information. Following completion of the information sheet, a rheumatologist conducted a comprehensive evaluation for FM based on the ACR criteria (Wolfe et al., 1990). Participants who met the ACR criteria and did not possess the exclusion criteria participated in the experimental sessions.

In one experimental session, participants were instructed to give their “normal effort” when completing a series of questionnaires and tests. In that session, the experimenter met the participant alone in the testing room, gave the participant a written copy of the experimental instructions, and read the instructions to the participant. Those instructions are listed below.

Medical problems, such as Fibromyalgia, are often poorly understood.

In this study, we are trying to help clarify some issues in the assessment of Fibromyalgia. You will complete a series of tests that
are designed to measure symptoms and problems that might be related to Fibromyalgia. Examples include problems related to pain, physical strength, daily functioning, mood, sleep, and memory.

In this session, it is important that you complete all of the tests and questionnaires carefully and honestly, in a manner that is normal for you.

The experimenter answered any questions the participant had regarding the instructions, handed a copy of the instructions to the participant in an envelope and left the room. The research assistant returned and reviewed the instructions of several questionnaires with the participant, answered any questions related to the questionnaires, and left the participant alone to complete the questionnaires. Those questionnaires included the Multidimensional Pain Inventory-1, Beck Depression Inventory-II, Pittsburgh Sleep Quality Index, and the Oswestry Disability Index-2.0. Once the participant completed the questionnaires the experimenter returned to the testing room, re-read the experimental instructions, answered any further questions regarding the instructions, and left the room. The research assistant returned and administered the Victoria Symptom Validity Test (VSVT). Once the VSVT was completed, instructions for the Personality Assessment Inventory (PAI) were reviewed, and the participant was left alone to complete that test. Upon completion of the PAI, the experimenter returned, pointed to the envelope containing the experimental instructions, reminded the participant to continue following the instructions, and left the room. The research assistant returned and administered the Test of Memory Malingering (TOMM). During the delay interval of the TOMM, the 21 Item Test was administered.

Participants were encouraged to take breaks prior to the administration of the Victoria Symptom Validity Test, the Test of Memory Malingering, and the 21 Item Test, because those tests required completion without interruption. Otherwise, participants were allowed to take breaks throughout the test session in order to stretch and alleviate pain.

In the other experimental session, participants were instructed to “exaggerate disability” due to fibromyalgia on the same measures. The introductory paragraph was the same for both experimental sessions. However, the “normal effort” instructions contained in
the second paragraph of the instructions described above were replaced with the instructions listed below in the “exaggerate disability” session.

**In this session, it is important that you complete all of the tests and questionnaires in a manner that would demonstrate high levels of pain and disability, and other problems associated with Fibromyalgia. Please exaggerate problems in a manner that would be convincing to others around you.**

In this session, the experimenter met each participant individually in the testing room, gave the participant a copy of the experimental instructions, and asked the participant to read the instructions while the experimenter left the room for a couple of minutes. The experimenter returned, read the instructions to the participant, and answered any questions the participant had regarding those instructions. The experimenter handed a copy of the instructions in an envelope to the participant and left the room. The remainder of this session was conducted in the same manner as the experimental session described above.

All participants were given a post-experimental interview (see Appendix B) at the end of the session involving the exaggeration instructions in order to assess: (a) whether the participant exaggerated, (b) whether they had any difficulties exaggerating, and (c) what strategies they used when exaggerating FM.

A health insurance inventory (see Appendix C) was given at the end of the second session in order to assess: (a) participants' understanding of their obligations in reporting their health status in a health insurance policy, and (b) how exaggerating pain and disability would affect that policy.

The medical screening assessment took approximately 15 minutes, and each experimental session took approximately 3 hours. Both experimental sessions were completed in a quiet room in the same building as the medical assessment. Research assistants were blind to the experimental conditions and the purpose of the study.

This is an analog study in which FM participants completed all experimental procedures under the two conditions described above. To counterbalance possible sequencing effects, half of the participants received the instructions for “normal” effort in the first experimental session and instructions to “exaggerate disability” in the second experimental session, while the remaining participants received the instructions in the reverse order.
Statistical Analyses

Statistical analyses were based on a within-subjects design. The independent variable, or within-subjects factor, is instructional set (normal effort and exaggerate disability). The dependent variables are test scores on the various measures (psychological, symptom validity, and personality). Participants were compared on all dependent measures to determine if there was a between-subject effect attributable to the sequence of instructional set. Repeated measures MANOVAs for psychological tests, Multidimensional Pain Inventory scales, Personality Assessment Inventory scales, and effort tests all revealed no significant differences in test scores attributable to order of experimental instructions. Therefore, the groups were combined into single “normal” and “exaggeration” conditions.

Descriptive statistics were used in the preliminary analyses to describe and summarize the variables. Type I error was controlled for by a Bonferroni correction to each omnibus MANOVA (alpha = .016). This correction was derived by dividing the overall alpha by the number of omnibus MANOVAs (.05 / 3). The results of each MANOVA were reported according to Wilk’s Lambda criterion. Univariate ANOVAs were used for follow-up comparisons for every measure following each MANOVA. In all cases, MANOVAs were assessed for violations of sphericity using Mauchly’s W test, and sphericity was satisfactory in all cases.

Results from the post-experimental questionnaire (as summarized below) indicated that 91% (n = 49) of the patients exaggerated disability on the psychological tests and questionnaires. Therefore, the remaining 9% of patients who did not exaggerate were excluded from the descriptive statistics and analyses evaluating performance on the psychological tests and the Personality Assessment Inventory. Results from the post-experimental questionnaire also indicated that the vast majority of patients (78%) did not exaggerate on the effort tests. Given that the exclusion of patients who did not exaggerate on the effort tests would have resulted in an insufficient sample size for analyses, the evaluation of effort test performance included the entire patient sample.
Results

Evaluation of Psychological Test Performance

Descriptive statistics for the psychological tests are provided in Table 16. Performance on the five psychological tests (BDI-II, MPI-I, FIQ, ODI-2, and the PSQI) was examined with a repeated measures MANOVA. There was a significant main effect for experimental instruction according to Wilk’s criterion, \( F(5, 44) = 24.35, p < .0001, \eta^2 = .735 \). Follow-up univariate ANOVAs for test scores across experimental conditions revealed significantly greater levels of depression (BDI-II), pain severity (MPI-I: Pain Severity scale), functional disability (ODI-2), sleep disturbance (PSQI), decreased health status (FIQ), and decreased general activity level (MPI-I) for the exaggeration condition versus the normal condition (see Table 16). In terms of effect size, or the magnitude of difference between experimental conditions, the effect sizes for all test score differences were in the large (FIQ, PSQI), to very large (BDI-II, MPI-I, ODI-2) range.

Table 16
Psychological test scores by experimental condition

<table>
<thead>
<tr>
<th>Test</th>
<th>Normal</th>
<th>Exaggeration</th>
<th>F</th>
<th>p value</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory</td>
<td>20.0 10.8</td>
<td>35.3 14.8</td>
<td>73.97</td>
<td>.0001</td>
<td>1.20</td>
</tr>
<tr>
<td>MPI (Pain Severity scale)</td>
<td>3.82 1.05</td>
<td>4.90 .88</td>
<td>60.30</td>
<td>.0001</td>
<td>1.12</td>
</tr>
<tr>
<td>FM Impact Questionnaire</td>
<td>60.44 16.2</td>
<td>73.18 15.44</td>
<td>36.60</td>
<td>.0001</td>
<td>.81</td>
</tr>
<tr>
<td>Oswestry Disability Index</td>
<td>36.79 15.96</td>
<td>56.36 17.76</td>
<td>96.55</td>
<td>.0001</td>
<td>1.16</td>
</tr>
<tr>
<td>Pittsburgh Sleep Quality Index</td>
<td>11.41 4.17</td>
<td>14.35 4.03</td>
<td>37.56</td>
<td>.0001</td>
<td>.72</td>
</tr>
</tbody>
</table>

Note. Sample size \( (N) = 49 \). Abbreviations: MPI = Multidimensional Pain Inventory; \( d \) = effect size, or the magnitude of difference between two groups in SD units.

The Multidimensional Pain Inventory (MPI) score included in the MANOVA examining the five psychological tests (see Table 17) was the Pain Severity scale score, which is only representative of one scale on that test. The MPI involves several scales, each scale measuring different aspects of pain-related issues. Therefore, a repeated measures MANOVA was conducted on all 13 of the MPI scales to more thoroughly examine effects of the experimental instruction on various pain-related issues. There was a main effect for experimental instruction according to Wilks’ criterion \( [F(13, 24) = 3.30, p<.005, \eta^2 = .641] \). Follow-up univariate ANOVAs for each scale scores across experimental conditions revealed significantly greater pain severity, interference, affective distress, support, and negative
responses from other, as well as significantly decreased life control and participation in household chores in the exaggeration condition, compared to the normal condition (see Table 17). The effect size between experimental conditions was large to very large for affective distress, life control, and pain severity. Otherwise, the effect sizes for all remaining significant scale score differences were in the medium to large range.

Table 17
Multidimensional Pain Inventory test scores by experimental condition

<table>
<thead>
<tr>
<th></th>
<th>Normal Mean</th>
<th>Normal SD</th>
<th>Exaggeration Mean</th>
<th>Exaggeration SD</th>
<th>F</th>
<th>p value</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain Impact</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain Severity</td>
<td>3.86</td>
<td>1.06</td>
<td>4.89</td>
<td>.92</td>
<td>38.54</td>
<td>.0001</td>
<td>1.04</td>
</tr>
<tr>
<td>Interference</td>
<td>3.87</td>
<td>1.33</td>
<td>4.67</td>
<td>1.19</td>
<td>16.92</td>
<td>.0001</td>
<td>.61</td>
</tr>
<tr>
<td>Life Control</td>
<td>3.61</td>
<td>1.19</td>
<td>2.54</td>
<td>1.48</td>
<td>18.16</td>
<td>.0001</td>
<td>.80</td>
</tr>
<tr>
<td>Affective Distress</td>
<td>3.32</td>
<td>1.0</td>
<td>4.25</td>
<td>.91</td>
<td>16.98</td>
<td>.0001</td>
<td>.97</td>
</tr>
<tr>
<td>Support</td>
<td>3.34</td>
<td>1.70</td>
<td>3.48</td>
<td>1.87</td>
<td>.68</td>
<td>.415</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Response by Significant Others</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>2.06</td>
<td>1.39</td>
<td>2.80</td>
<td>1.74</td>
<td>7.55</td>
<td>.009</td>
<td>.47</td>
</tr>
<tr>
<td>Solicitous</td>
<td>2.86</td>
<td>1.64</td>
<td>2.74</td>
<td>1.83</td>
<td>.39</td>
<td>.536</td>
<td>.07</td>
</tr>
<tr>
<td>Distracting</td>
<td>2.02</td>
<td>1.29</td>
<td>2.03</td>
<td>1.47</td>
<td>.006</td>
<td>.939</td>
<td>.01</td>
</tr>
<tr>
<td><strong>Participation in Daily Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Chores</td>
<td>4.10</td>
<td>1.71</td>
<td>3.35</td>
<td>1.72</td>
<td>14.35</td>
<td>.001</td>
<td>.44</td>
</tr>
<tr>
<td>Outdoor Work</td>
<td>1.35</td>
<td>1.18</td>
<td>1.06</td>
<td>1.30</td>
<td>2.87</td>
<td>.099</td>
<td>.19</td>
</tr>
<tr>
<td>Activities Away From Home</td>
<td>2.72</td>
<td>.89</td>
<td>2.82</td>
<td>1.27</td>
<td>.29</td>
<td>.592</td>
<td>.09</td>
</tr>
<tr>
<td>Social Activities</td>
<td>2.56</td>
<td>1.16</td>
<td>2.31</td>
<td>1.47</td>
<td>3.16</td>
<td>.084</td>
<td>.19</td>
</tr>
<tr>
<td>General Activity Level</td>
<td>2.83</td>
<td>1.21</td>
<td>2.59</td>
<td>1.39</td>
<td>3.95</td>
<td>.054</td>
<td>.18</td>
</tr>
</tbody>
</table>

Note. The sample size (N) = 37, because 12 subjects did not complete some of the scales. Abbreviations: d = effect size, or the magnitude of difference between two groups in SD units.

Evaluation of Personality Assessment Inventory Performance

Descriptive statistics for the Personality Assessment Inventory validity and clinical scale scores are provided in Table 18. Performance on the seven validity scales and indexes, and the 11 clinical scales, were examined with a repeated measures MANOVA. A main effect for experimental condition was significant according to Wilk's criterion, F (18, 31) = 3.15, p < .002, η² = .646. Follow-up univariate ANOVAs for test scores across experimental conditions revealed significantly higher scores in the exaggeration condition for all of the validity and index scale scores, except for the Infrequency scale. Scores were significantly lower in the exaggeration condition for the Positive Impression Management scale. In addition, significantly higher scores were present in the exaggeration condition for all of the clinical scales (see Table 18).
Table 18

Personality Assessment Inventory scale and index scores by experimental condition

<table>
<thead>
<tr>
<th>Validity Scales &amp; Indexes</th>
<th>Normal</th>
<th>Exaggeration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Inconsistency</td>
<td>51.20</td>
<td>8.72</td>
</tr>
<tr>
<td>Infrequency</td>
<td>49.43</td>
<td>7.72</td>
</tr>
<tr>
<td>Negative Impression</td>
<td>58.37</td>
<td>11.38</td>
</tr>
<tr>
<td>Positive Impression</td>
<td>48.04</td>
<td>10.63</td>
</tr>
<tr>
<td>Defensiveness Index</td>
<td>51.51</td>
<td>10.28</td>
</tr>
<tr>
<td>Malingering Index</td>
<td>57.59</td>
<td>13.86</td>
</tr>
<tr>
<td>Rogers Discriminant Function</td>
<td>49.22</td>
<td>9.49</td>
</tr>
<tr>
<td>Clinical Scales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic Complaints</td>
<td>73.43</td>
<td>13.33</td>
</tr>
<tr>
<td>Anxiety</td>
<td>61.57</td>
<td>12.58</td>
</tr>
<tr>
<td>Anxiety Related Disorders</td>
<td>60.16</td>
<td>14.35</td>
</tr>
<tr>
<td>Depression</td>
<td>66.69</td>
<td>12.06</td>
</tr>
<tr>
<td>Mania</td>
<td>50.49</td>
<td>10.89</td>
</tr>
<tr>
<td>Paranoia</td>
<td>48.90</td>
<td>8.51</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>55.02</td>
<td>11.57</td>
</tr>
<tr>
<td>Borderline Features</td>
<td>53.98</td>
<td>11.47</td>
</tr>
<tr>
<td>Antisocial Features</td>
<td>46.10</td>
<td>7.89</td>
</tr>
<tr>
<td>Alcohol Problems</td>
<td>47.10</td>
<td>6.76</td>
</tr>
<tr>
<td>Drug Problems</td>
<td>51.10</td>
<td>9.82</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. All test scores reported are T-scores. d = effect size, or the magnitude of difference between two groups in SD units.

Evaluation of Effort Test Performance

Descriptive statistics for the effort tests are provided in Table 19. Performances on the three effort tests (VSVT, TOMM, 21 Item Test) and the scores primarily used in evaluating effort were examined with a repeated measures MANOVA. The primary scores included the Hard, Easy, and Overall (VSVT), Trial 2 and Retention Trial (TOMM), and the Forced Choice (21 Item Test) scores. The main effect for experimental condition was not significant according to Wilk's criterion, $F(5, 49) = 1.61, p < .176, \eta^2 = .141$. Exploratory univariate ANOVAs for primary test scores across experimental conditions revealed significantly lower VSVT total scores and Hard Item scores in the exaggeration condition (see Table 19).

Performances on the three effort tests with the secondary or remaining scores were also examined with a repeated measures MANOVA. The secondary scores included the Easy and Hard Response Time (VSVT), Trial 1 (TOMM), and the Inconsistencies and Great Consecutive Misses (21 Item Test) scores. Again, the main effect for experimental condition
was not significant according to Wilk’s criterion, \( F(5, 49) = 1.2, p < .323, \eta^2 = .109. \)

Exploratory univariate ANOVAs for primary test scores across experimental conditions revealed significantly lower TOMM Trial 1 scores in the exaggeration condition.

Table 19
Effort test scores by experimental condition for all patients (N=54)

<table>
<thead>
<tr>
<th>Test</th>
<th>Normal</th>
<th>Exaggeration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>VSVT - Overall correct</td>
<td>46.2</td>
<td>2.8</td>
</tr>
<tr>
<td>- Easy Items Correct</td>
<td>23.7</td>
<td>.8</td>
</tr>
<tr>
<td>- Hard Items Correct</td>
<td>22.5</td>
<td>2.3</td>
</tr>
<tr>
<td>- Easy Response Time</td>
<td>1.3</td>
<td>.8</td>
</tr>
<tr>
<td>- Hard Response Time</td>
<td>2.2</td>
<td>1.3</td>
</tr>
<tr>
<td>TOMM Trial 1</td>
<td>48.8</td>
<td>1.9</td>
</tr>
<tr>
<td>TOMM Trial 2</td>
<td>49.8</td>
<td>.5</td>
</tr>
<tr>
<td>TOMM Retention</td>
<td>49.6</td>
<td>.9</td>
</tr>
<tr>
<td>21 Item Test – Forced Choice</td>
<td>17.4</td>
<td>2.2</td>
</tr>
<tr>
<td>21 Item Test – Inconsistencies</td>
<td>.2</td>
<td>.5</td>
</tr>
<tr>
<td>21 Item Test – GCM</td>
<td>1.6</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Note. d = effect size, or the magnitude of difference between two groups in SD units. Abbreviations: VSVT = Victoria Symptom Validity Test; TOMM = Test of Memory Malingering; GCM = Greatest Consecutive Misses.

Patients obtained high scores on all measures of the VSVT, all three trials of the TOMM, and both measures of the 21 Item Test in the normal and exaggeration conditions.

The VSVT manual indicates that total items correct scores between 18-29, and easy and hard items correct between 8-15 are in the “questionable” range, indicating the possibility of feigning impairment (Slick et al., 1997). In addition, easy items response latencies above 2.84 seconds and hard item response latencies above 5.50 seconds also indicate the possibility of feigning impairment. A total of 89% of patients in the normal condition, and 69% of patients in the exaggeration condition scored above the cutoffs on these measures, indicating valid scores on the VSVT. The remaining 11% of patients in the normal condition scored below the cutoff, and in the “questionable” range. Out of those patients, one scored below the cutoff on the hard items correct, five scored below the cutoff on the easy item response latencies, and one scored below the cutoff on the hard item response latencies. In the exaggeration condition, 31% of patients scored below the cutoff, and in the “questionable” range. Of those patients, three scored below the cutoff on the total items correct, one scored below the cutoff on the easy items correct, nine scored below the cutoff
on the hard items correct, two scored below the cutoff on the easy response latencies, and one scored below the cutoff on the hard items response latencies.

The TOMM manual indicates that any score lower than 45 on Trial 2 or on the Retention Trial may indicate poor effort (Tombaugh, 1996). All patients in the normal condition scored above the cutoffs, indicating perfect specificity. However, only 11% of patients in the exaggeration condition scored below the cutoffs, suggesting extremely low sensitivity.

Finally, the manual for the 21 Item Test indicates that a score of 12 or below on the Forced Choice section is suggestive of poor effort. A total of 96% of patients in the normal condition and 93% of patients in the exaggeration condition scored above this cutoff. The manual also contains two decision rules that can be applied to the scores. One decision rule indicates that a score below 9 on the Forced Choice section; or greater than 4 on the Inconsistency score (the number of words correctly recalled during free recall that were subsequently missed on the forced-choice section); or greater than 6 on the Greatest Consecutive Misses score (the greatest number of consecutive misses on the forced-choice section) suggests poor effort. According to the first decision rule, none of the patients in the control condition, and 2% of patients in the exaggeration condition met this rule. The second decision rule was similar to the first rule, but involved a score of below 12 on the Forced Choice section, greater than 2 on the Inconsistency score, or greater than 5 on the Greatest Consecutive Misses score. According to the second decision rule, 2% of patients in the control condition and 6% of patients in the exaggeration condition met this rule.

Following the exaggeration condition, patients indicated whether they had exaggerated on the effort tests in a post-experimental questionnaire. Of the total sample, only 12 patients indicated that they had exaggerated on the effort tests. These individuals were examined separately. The main effect for experimental condition with the primary scores was not significant according to Wilk’s criterion \[F(5, 7) = 1.96, p < .203, \eta^2 = .583\]. Exploratory univariate ANOVAs for test scores across experimental conditions revealed significantly lower VSVT total scores and Hard Item scores, in the exaggeration condition (see Table 20). The main effect for experimental condition with the secondary scores was also not significant according to Wilk’s criterion \[F(5, 7) = .835, p < .564, \eta^2 = .374\].
Table 20

Effort test scores by experimental condition only for exaggerators (n=12)

<table>
<thead>
<tr>
<th>Test</th>
<th>Normal</th>
<th></th>
<th>Exaggeration</th>
<th></th>
<th>F</th>
<th>p value</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSVT - Overall correct</td>
<td>46.92</td>
<td>1.31</td>
<td>38.67</td>
<td>7.71</td>
<td>11.31</td>
<td>.006</td>
<td>1.82</td>
</tr>
<tr>
<td>- Easy Items Correct</td>
<td>23.92</td>
<td>.29</td>
<td>23.00</td>
<td>1.91</td>
<td>2.58</td>
<td>.136</td>
<td>.82</td>
</tr>
<tr>
<td>- Hard Items Correct</td>
<td>23.00</td>
<td>1.21</td>
<td>15.67</td>
<td>6.46</td>
<td>12.48</td>
<td>.005</td>
<td>1.90</td>
</tr>
<tr>
<td>- Easy Response Time</td>
<td>1.38</td>
<td>.63</td>
<td>1.47</td>
<td>.57</td>
<td>.58</td>
<td>.463</td>
<td>.17</td>
</tr>
<tr>
<td>- Hard Response Time</td>
<td>2.59</td>
<td>1.27</td>
<td>2.68</td>
<td>1.51</td>
<td>.15</td>
<td>.701</td>
<td>.07</td>
</tr>
<tr>
<td>TOMM Trial 1</td>
<td>48.67</td>
<td>1.87</td>
<td>46.17</td>
<td>6.26</td>
<td>1.51</td>
<td>.244</td>
<td>.62</td>
</tr>
<tr>
<td>TOMM Trial 2</td>
<td>50.00</td>
<td>.00</td>
<td>46.67</td>
<td>6.21</td>
<td>3.45</td>
<td>.090</td>
<td>1.06</td>
</tr>
<tr>
<td>TOMM Retention</td>
<td>49.67</td>
<td>.49</td>
<td>46.67</td>
<td>6.01</td>
<td>3.06</td>
<td>.108</td>
<td>.92</td>
</tr>
<tr>
<td>21 Item Test – Forced Choice</td>
<td>17.67</td>
<td>1.83</td>
<td>16.83</td>
<td>2.52</td>
<td>1.28</td>
<td>.282</td>
<td>.41</td>
</tr>
<tr>
<td>21 Item Test – Inconsistencies</td>
<td>.0</td>
<td>.0</td>
<td>.1</td>
<td>.3</td>
<td>1.00</td>
<td>.339</td>
<td>.67</td>
</tr>
<tr>
<td>21 Item Test - GCM</td>
<td>1.6</td>
<td>.9</td>
<td>1.4</td>
<td>.7</td>
<td>.31</td>
<td>.586</td>
<td>.25</td>
</tr>
</tbody>
</table>

Note. d = effect size, or the magnitude of difference between two groups in SD units. Abbreviations: VSVT = Victoria Symptom Validity Test; TOMM = Test of Memory Malingering; GCM = Greatest Consecutive Misses.

Of the sub-sample of 12 patients indicating that they exaggerated on the effort tests, a 92% specificity rate (correct identification of non-exaggerators) and a 58% sensitivity rate (correct identification of exaggerators) was obtained for the VSVT. There was a perfect specificity rate and a 33% sensitivity rate for those patients on the TOMM. Finally, there was a perfect specificity rate and a 2% sensitivity rate for those patients on the Forced Choice section of the 21 Item Test.

Examination of Frequency Distributions

Frequency score distributions were examined to select cutoff scores that would correctly classify as many patients as possible according to experimental condition. Cutoff scores, based on examination of group frequency distributions, were chosen to minimize false positives. Cutoff scores and classification rates for the psychological tests appear in Table 21. Multidimensional Pain Inventory scales with an effect size (d) greater than .60 (medium effect size) also appear in Table 21. The Life Control scale was excluded from this process because no cutoff score with a low false positive rate could be derived.

Several cutoff scores are presented in Table 21 illustrating different hit and false positive rates. The hit rates for correctly identifying exaggeration ranged from 10 – 63%, with false positive rates ranging from 0 – 18%. For nearly every cutoff score, the majority of exaggerators were not detected. The cutoff scores with the highest hit rates while keeping false positives below 10% are bolded in Table 21.
Table 21
Cutoffs, accuracy, and false positive rates for psychological test scores

<table>
<thead>
<tr>
<th>Test</th>
<th>Cutoff Score</th>
<th>Normal (%)</th>
<th>Exaggeration (%)</th>
<th>Total Classified (%)</th>
<th>False Positives (%)</th>
<th>False Negatives (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI - II</td>
<td>&gt;28&lt;sup&gt;1&lt;/sup&gt;</td>
<td>82</td>
<td>63</td>
<td>72</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>&gt;36&lt;sup&gt;1&lt;/sup&gt;</td>
<td>92</td>
<td>43</td>
<td>67</td>
<td>8</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>&gt;37</td>
<td>96</td>
<td>41</td>
<td>68</td>
<td>4</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>100</td>
<td>18</td>
<td>59</td>
<td>0</td>
<td>82</td>
</tr>
<tr>
<td>FIQ</td>
<td>&gt;78.9</td>
<td>94</td>
<td>41</td>
<td>67</td>
<td>6</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>&gt;92.0</td>
<td>98</td>
<td>12</td>
<td>55</td>
<td>2</td>
<td>88</td>
</tr>
<tr>
<td>ODI - 2.0</td>
<td>&gt;62</td>
<td>92</td>
<td>35</td>
<td>63</td>
<td>8</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>&gt;74</td>
<td>100</td>
<td>16</td>
<td>58</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>PSQI</td>
<td>&gt;17</td>
<td>92</td>
<td>24</td>
<td>58</td>
<td>8</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>&gt;19</td>
<td>98</td>
<td>14</td>
<td>56</td>
<td>2</td>
<td>86</td>
</tr>
<tr>
<td>Pain Severity (MPI)</td>
<td>&gt;4.70</td>
<td>88</td>
<td>59</td>
<td>73</td>
<td>12</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>&gt;5.00</td>
<td>92</td>
<td>37</td>
<td>64</td>
<td>8</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>&gt;5.67</td>
<td>98</td>
<td>18</td>
<td>58</td>
<td>2</td>
<td>82</td>
</tr>
<tr>
<td>Interference (MPI)</td>
<td>&gt;5.80</td>
<td>96</td>
<td>14</td>
<td>55</td>
<td>4</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>&gt;5.90</td>
<td>98</td>
<td>10</td>
<td>54</td>
<td>2</td>
<td>90</td>
</tr>
<tr>
<td>Affective Distress (MPI)</td>
<td>&gt;4.33</td>
<td>92</td>
<td>45</td>
<td>68</td>
<td>8</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>&gt;4.67</td>
<td>94</td>
<td>29</td>
<td>61</td>
<td>6</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>&gt;5.30</td>
<td>98</td>
<td>16</td>
<td>57</td>
<td>2</td>
<td>84</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. <sup>1</sup>Cutoff scores between 29 and 36 represent severe depression (Beck et al., 1996).

Abbreviations: BDI-II = Beck Depression Inventory- version 2; MPI = Multidimensional Pain Inventory – version 1.0; FIQ = Fibromyalgia Impact Questionnaire; ODI-2.0 = Oswestry Disability Index-version 2.0; PSQI = Pittsburgh Sleep Quality Index. Cutoff scores with the highest hit rates while keeping false positive rates below 10% are bolded.

It has been suggested that multiple test score combinations may improve the identification rate of exaggeration over individual tests (Bianchini et al., 2005; Iverson et al., 1996; Meyers et al., 2002). Hence, a series of multiple cutoff score combinations were examined for the psychological tests to determine whether they improved classification rates. Combinations of the Beck Depression Inventory - II (BDI), FM Impact Questionnaire (FIQ), Oswestry Disability Index – 2 (ODI), and Pittsburgh Sleep Quality Index (PSQI) were examined. Some scales from the Multidimensional Pain Inventory - 1 (MPI) were also examined because they had a large effect size in the ANOVA results, including the Pain Severity, Interference, and Affective Distress scale. The cutoff scores chosen for each scale were associated with the highest correct classification rates and minimal false positive rates (less than 10%). Cutoffs for the psychological tests were as follows: BDI (37), FIQ (78.9),...
OAI (62), PSQI (17), MPI: Pain Severity (4.7), MPI: Interference (5.8), and MPI: Affective Distress (4.33).

The first test score combination involved three psychological tests. As seen in Table 22, scoring above the cutoff on one or more tests resulted in a 57% hit rate with 14% false positives. Having two or more high scores correctly identified 37% of exaggerators with 4% false positives.

Table 22
Cumulative frequency (%) of score combinations for the BDI, FIQ, and ODI

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>0</td>
<td>86</td>
<td>86</td>
<td>na</td>
<td>43</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>96</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>98</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>100</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: BDI=37; FIQ=78.9; ODI=62. There are slight deviations due to rounding.

The second test score combination involved three psychological tests and the MPI Pain Severity scale. As seen in Table 23, scoring above the cutoff on one or more tests resulted in a 69% hit rate with 22% false positives. Having two or more high scores correctly identified 53% of exaggerators with a 4% false positive rate.

Table 23
Cumulative frequency (%) of score combinations for the BDI, FIQ, ODI, and MPI-Pain Severity scale

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>0</td>
<td>78</td>
<td>78</td>
<td>na</td>
<td>31</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>96</td>
<td>22</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>98</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>100</td>
<td>2</td>
<td>18</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: BDI=37; FIQ=78.9; ODI=62; Pain Severity=4.7. There are slight deviations due to rounding.

The third test score combination involved all of the psychological tests and the MPI Pain Severity scale. As seen in Table 24, scoring above the cutoff on one or more tests
resulted in an 80% hit rate with 29% false positives. Having two or more high scores correctly identified 61% of exaggerators with 12% false positives.

Table 24
Cumulative frequency (%) of score combinations for the BDI, FIQ, ODI, PSQI, and MPI-
Pain Severity, Interference, and Affective Distress scales

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>0</td>
<td>71</td>
<td>71</td>
<td>na</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>88</td>
<td>29</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>94</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>96</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>98</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>100</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: BDI=37; FIQ=78.9; ODI=62; PSQI=17; Pain Severity=4.7; Interference=5.8; Affective Distress=4.33. There are slight deviations due to rounding.

The final score combination involved the three MPI scales. As seen in Table 25, the majority (80%) of patients in the control condition did not reach the cutoff on any of these scores. In contrast, the majority of patients in the exaggeration condition (i.e., 73%) obtained one or more scores above the cutoffs. If the criterion for exaggeration was two or more scores, the hit rate would be 33% and the false positive rate would be 4%.

Table 25
Cumulative frequency (%) of scores combinations for the MPI-Pain Severity, Interference, and Affective Distress scales

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>P</td>
<td>p</td>
</tr>
<tr>
<td>0</td>
<td>80</td>
<td>80</td>
<td>na</td>
<td>27</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>96</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>100</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>12</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: Pain Severity=4.7; Interference=5.8; Affective Distress=4.33. There are slight deviations due to rounding.

Cutoff scores and respective accuracy rates for classification into experimental condition from the Personality Assessment Inventory scale scores appear in Table 26. As
seen in the total classification column, the percentages of patients correctly classified was, essentially, no better than chance (flipping a coin) for most measures.

Table 26
Cutoffs, accuracy, and false positive rates for Personality Assessment Inventory scales and indexes

<table>
<thead>
<tr>
<th>Validity Scales &amp; Indexes</th>
<th>Cutoff Score</th>
<th>Normal (%)</th>
<th>Exaggeration (%)</th>
<th>Total Classified (%)</th>
<th>False Positives (%)</th>
<th>False Negatives (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inconsistency</td>
<td>&gt;64</td>
<td>98</td>
<td>18</td>
<td>58</td>
<td>2</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>&gt;73&lt;sup&gt;1&lt;/sup&gt;</td>
<td>90</td>
<td>24</td>
<td>57</td>
<td>10</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>&gt;77&lt;sup&gt;2&lt;/sup&gt;</td>
<td>94</td>
<td>20</td>
<td>57</td>
<td>6</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>&gt;81</td>
<td>98</td>
<td>18</td>
<td>58</td>
<td>2</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>&gt;84&lt;sup&gt;4&lt;/sup&gt;</td>
<td>98</td>
<td>12</td>
<td>55</td>
<td>2</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>&gt;92&lt;sup&gt;1&lt;/sup&gt;</td>
<td>100</td>
<td>8</td>
<td>54</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>Negative Impression</td>
<td>&gt;61</td>
<td>96</td>
<td>2</td>
<td>49</td>
<td>4</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>&gt;64</td>
<td>98</td>
<td>6</td>
<td>52</td>
<td>2</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>&gt;71</td>
<td>92</td>
<td>29</td>
<td>60</td>
<td>8</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>&gt;84&lt;sup&gt;3&lt;/sup&gt;</td>
<td>98</td>
<td>8</td>
<td>50</td>
<td>2</td>
<td>92</td>
</tr>
<tr>
<td>Positive Impression</td>
<td>&gt;64</td>
<td>96</td>
<td>20</td>
<td>58</td>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>&gt;70&lt;sup&gt;4&lt;/sup&gt;</td>
<td>100</td>
<td>8</td>
<td>54</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>Defensiveness Index</td>
<td>&gt;64</td>
<td>98</td>
<td>6</td>
<td>52</td>
<td>2</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>&gt;71</td>
<td>92</td>
<td>29</td>
<td>60</td>
<td>8</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>&gt;84&lt;sup&gt;3&lt;/sup&gt;</td>
<td>98</td>
<td>8</td>
<td>50</td>
<td>2</td>
<td>92</td>
</tr>
<tr>
<td>Malingering Index</td>
<td>&gt;64</td>
<td>96</td>
<td>20</td>
<td>58</td>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>&gt;70&lt;sup&gt;4&lt;/sup&gt;</td>
<td>100</td>
<td>8</td>
<td>54</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>Rogers Discriminant</td>
<td>&gt;70&lt;sup&gt;4&lt;/sup&gt;</td>
<td>100</td>
<td>8</td>
<td>54</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>Clinical Scales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic Complaints</td>
<td>&gt;93</td>
<td>92</td>
<td>20</td>
<td>56</td>
<td>8</td>
<td>80</td>
</tr>
<tr>
<td>Anxiety</td>
<td>&gt;81</td>
<td>94</td>
<td>22</td>
<td>58</td>
<td>6</td>
<td>78</td>
</tr>
<tr>
<td>Anxiety Related Disorders</td>
<td>&gt;78</td>
<td>94</td>
<td>31</td>
<td>62</td>
<td>6</td>
<td>69</td>
</tr>
<tr>
<td>Depression</td>
<td>&gt;86</td>
<td>94</td>
<td>33</td>
<td>63</td>
<td>6</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>&gt;93</td>
<td>100</td>
<td>18</td>
<td>59</td>
<td>0</td>
<td>82</td>
</tr>
<tr>
<td>Mania</td>
<td>&gt;70</td>
<td>96</td>
<td>16</td>
<td>56</td>
<td>4</td>
<td>84</td>
</tr>
<tr>
<td>Paranoia</td>
<td>&gt;74</td>
<td>100</td>
<td>10</td>
<td>55</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>&gt;73</td>
<td>96</td>
<td>22</td>
<td>59</td>
<td>4</td>
<td>78</td>
</tr>
<tr>
<td>Borderline Features</td>
<td>&gt;72</td>
<td>96</td>
<td>27</td>
<td>61</td>
<td>4</td>
<td>73</td>
</tr>
<tr>
<td>Antisocial Features</td>
<td>&gt;60</td>
<td>96</td>
<td>12</td>
<td>54</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>Alcohol Problems</td>
<td>&gt;65</td>
<td>98</td>
<td>12</td>
<td>55</td>
<td>2</td>
<td>88</td>
</tr>
<tr>
<td>Drug Problems</td>
<td>&gt;66</td>
<td>94</td>
<td>12</td>
<td>53</td>
<td>6</td>
<td>88</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Infrequency scale cutoff scores were not calculated due to nonsignificant follow-up results. <sup>1</sup>Negative Impression Management cutoff scores >73 indicative of simulators for normal population, >84 indicative of simulators for clinical population, >92 considered indicative of malingering (Morey, 1991).<sup>2</sup>Negative Impression Management cutoff scores >77 moderately successful at detecting simulated psychological disorders (Rogers et al., 1996). <sup>3</sup>Malingering Index cutoff scores >84 indicative of malingering (Liljequist et al., 1998). <sup>4</sup>Rogers Discriminant Function cutoff scores >70 indicative of malingering (Poythress et al. (2001)).
The scores derived from the Personality Assessment Inventory (PAI) scales were not useful for identifying exaggeration. For most of the individual scale scores, fewer than 25% of the exaggerating patients were correctly identified.

A series of multiple cutoff score combinations were also examined for some of the PAI scales to determine whether they improved the effectiveness in predicting exaggeration over individual tests. In terms of validity scales and indexes, the NIM, MAL, and RDF were chosen because they were specifically intended to detect exaggeration and malingering. In terms of clinical scales, the Depression, Anxiety, Anxiety Related Disorders, Somatic Complaints, Schizophrenia, and Borderline Features were also selected because they were the only clinical scales with a large effect sizes in the ANOVA results. The cutoff scores chosen for each scale were associated with the highest correct classification rate and minimal false positive rates (less than 10%). Those cutoffs were as follows: NIM (81), MAL (71), RDF (70), Depression (86), Anxiety (81), Anxiety Related Disorders (78), Somatic Complaints (93), Schizophrenia (73), and Borderline Features (72).

Two test score combinations involving only the validity scales were examined, and the results are reported in Tables 27 and 28. As seen in both tables, the vast majority (90%) of patients in the control condition did not reach the cutoff on any of the validity scales, and that the use of one validity scale resulted in 10% of false positives. These tables also indicate that applying one or more validity scales correctly identifies 37-39% of patients who are exaggerating.

Table 27

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>p</td>
<td>cp</td>
</tr>
<tr>
<td>0</td>
<td>90</td>
<td>90</td>
<td>na</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>100</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td>12</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: NIM=81; MAL=71; RDF=70. There are slight deviations due to rounding.
Table 28
Cumulative frequency (%) of score combinations for the NIM and MAL

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>p</td>
<td>cp</td>
</tr>
<tr>
<td>0</td>
<td>90</td>
<td>90</td>
<td>na</td>
<td>63</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>100</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: NIM=81; MAL=71. There are slight deviations due to rounding.

Two test score combinations involving only the clinical scales were examined, and the results are reported in Tables 29 and 30. Table 29 involves all of the clinical scales with a large effect size from the ANOVA results, while Table 30 involves all of the clinical scales associated with symptoms of FM commonly reported in the medical literature. As seen in both tables, the vast majority (84%) of patients in the control condition did not reach the cutoff on any of these clinical scales. Requiring two or more scores above the cutoff results in 33-39% hit rates and 6-8% false positive rates.

Table 29
Cumulative frequency (%) of score combinations for the Depression, Anxiety, Anxiety Related, Somatic Complaints, Schizophrenia, and Borderline Features scales

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal</th>
<th>FP</th>
<th>Exaggeration</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>cp</td>
<td>p</td>
<td>cp</td>
</tr>
<tr>
<td>0</td>
<td>84</td>
<td>84</td>
<td>na</td>
<td>49</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>92</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>94</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>96</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>100</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: Depression=86; Anxiety=81; Anxiety Related=78; Somatic Complaints=93; Schizophrenia=73; Borderline Features=72. There are slight deviations due to rounding.
Table 30
Cumulative frequency (%) of score combinations for the Depression, Anxiety, Anxiety Related, and Somatic Complaints scales

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal FP</th>
<th>Exaggeration Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p cp p cp</td>
<td>p cp p cp</td>
</tr>
<tr>
<td>0</td>
<td>84 84 na 55</td>
<td>55 na</td>
</tr>
<tr>
<td>1</td>
<td>10 94 16 12</td>
<td>67 80 33</td>
</tr>
<tr>
<td>2</td>
<td>4 98 6 12</td>
<td>80 33</td>
</tr>
<tr>
<td>3</td>
<td>- - 2 12</td>
<td>92 20</td>
</tr>
<tr>
<td>4</td>
<td>2 100 2 8</td>
<td>100 8</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: Depression=86; Anxiety=81; Anxiety Related=78; Somatic Complaints=93. There are slight deviations due to rounding.

Finally, Table 31 involved a combination of two validity scales with all of the clinical scales that had a large effect size. As seen in this table, the majority (78%) of patients in the control condition did not reach the cutoff on any of the scales. Applying a criteria of two or more failures would result in a 45% hit rate with a 10% false positive rate.

Table 31
Cumulative frequency (%) of score combinations for the NIM, MAL, and selected clinical scales

<table>
<thead>
<tr>
<th>Number of Failed Scores</th>
<th>Normal FP</th>
<th>Exaggeration Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p cp p cp</td>
<td>p cp p cp</td>
</tr>
<tr>
<td>0</td>
<td>78 78 na 43</td>
<td>43 na</td>
</tr>
<tr>
<td>1</td>
<td>12 90 22 12</td>
<td>55 57</td>
</tr>
<tr>
<td>2</td>
<td>4 94 10 10</td>
<td>65 45</td>
</tr>
<tr>
<td>3</td>
<td>2 96 6 10</td>
<td>76 35</td>
</tr>
<tr>
<td>4</td>
<td>2 98 4 6</td>
<td>82 24</td>
</tr>
<tr>
<td>5</td>
<td>2 100 2 6</td>
<td>88 18</td>
</tr>
<tr>
<td>6</td>
<td>- - 0 6</td>
<td>94 12</td>
</tr>
<tr>
<td>7</td>
<td>- - 0 2</td>
<td>96 6</td>
</tr>
<tr>
<td>8</td>
<td>- - 0 4</td>
<td>100 4</td>
</tr>
</tbody>
</table>

Note. Sample size (N) = 49. Abbreviations: FP = false positives; p = percentage; cp = cumulative percentage; and na = not applicable. Cutoff scores are as follows: NIM=81; MAL=71; Depression=86; Anxiety=81; Anxiety Related=78; Somatic Complaints=93; Schizophrenia=73; Borderline Features=72. There are slight deviations due to rounding.

Post-Experimental Questionnaire
The post-experimental interview was designed to assess: (a) whether the participant exaggerated, (b) whether they had any difficulties exaggerating, and (c) what strategies they
used when exaggerating FM. Descriptive statistics were used to examine patient responses in that interview. Results revealed that 91% of patients indicated that they tried to exaggerate disability during the exaggeration condition. When asked if they had problems exaggerating disability overall, 52% of the patients acknowledged having problems, 22% acknowledged having problems on some measures, while 22% acknowledged no problems exaggerating disability. When asked what problems they had in exaggerating overall, those with problems indicated the following: 22% did not feel comfortable/felt it was wrong to exaggerate; 15% try to cope with pain (hence, do not typically exaggerate); 13% did not know how to exaggerate; 9% were already feeling bad and did not have to exaggerate; and 7% stated they had difficulty remembering to exaggerate.

In terms of exaggeration strategies for the psychological questionnaires, 54% rated their symptoms higher than was true, 26% thought of their worst pain days, 7% took the highest rating possible, 6% stated they were already feeling bad and did not have to exaggerate, 4% did not exaggerate, 4% focused on their pain (as opposed to coping with it), and 2% did not know. In terms of exaggeration strategies for the effort tests, 44% indicated that they did not exaggerate, 11% did not know how to exaggerate, 9% indicated they did not experience memory problems as part of FM, 7% purposely answered incorrectly, 4% could not exaggerate with someone present, 2% responded more slowly than normal, 2% thought the tests were so difficult that they did not have to exaggerate, 2% did not focus on answering correctly, and 2% did not know. Finally, patients rated how well they think they exaggerated disability on a scale of 1 to 10 for both the psychological questionnaires, and the effort tests. For the psychological questionnaires, 47% of patients gave a rating of 8 or greater, 32% gave a rating of 5 to 7.5, and 19% gave a rating below 5. For the effort tests, 8% of patients gave a rating of 8 or greater, 14% gave a rating of 5 to 7.5, and 80% gave a rating below 5.

Health Insurance Inventory

The health insurance inventory was designed to assess (a) patients' understanding of their obligations in reporting their health status in a health insurance policy, and (b) how exaggerating pain and disability would affect that policy. Descriptive statistics were used to examine patient responses on this inventory. Results revealed that 93% of patients have held a life, disability, or health insurance policy during their lifetime, 54% of patients have filed a claim related to a life, disability or health insurance policy, and that a claim was in progress.
for 24% of patients. In terms of health related legal claims, 74% of patients indicated never having filed a claim. FM was included in legal claims for 13% of patients, and all legal claims had been resolved at the time of the study. In terms of worker’s compensation claims due to medical problems, 76% of patients indicated never having filed a claim, claims were in progress for 6% of patients, and FM was included in 4% of worker’s compensation claims. Therefore, health insurance or worker’s compensation claims were in progress for 30% of the total patient sample during the time of the study, and FM was included in only 4% of those claims.

In terms of reporting past health-related information to the insurance agent, 76% of patients felt they were obligated to report all previous illnesses and health habits, 12% felt they were obligated to report only permanent, life threatening, or ongoing illnesses, 11% did not know their obligations, while 2% felt they were not obligated to report anything. In terms of reporting changes in health status to the insurance agent, 41% of patients felt they should report all changes, 20% did not know their obligations, 17% felt they were not obligated to report any changes, 13% felt they should only report surgeries, life-threatening, or permanent illnesses, and 9% felt that only changes that interfered with work ability should be reported.

Patients were also asked what would happen to their policy coverage if they either forgot to report changes in health, or exaggerated their medical symptoms. Patient responses are summarized in Table 32.

Table 32

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Unreported (%)</th>
<th>Exaggerated (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy cancellation/null/void</td>
<td>48</td>
<td>39</td>
</tr>
<tr>
<td>Denied partial or full policy coverage</td>
<td>26</td>
<td>30</td>
</tr>
<tr>
<td>Nothing would happen</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Do not know</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Incur penalty</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Withhold payment</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Receive more coverage</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>
Discussion

Determining the nature and extent of disability, and the accuracy of symptom reporting, is necessary for proper adjudication within the tort law system and proper insurance processing of disability claims. Issues of exaggeration, deliberate misrepresentation of disability level, and malingering add to the complexity of the claims process, especially for controversial disorders for which there is no obvious cause, like FM. It is therefore essential to explore methods that could assist in determining the credibility of those diagnosed with FM. The purpose of Study 2 was to examine the phenomenology of exaggeration and the deliberate portrayal of excessive disability in patients with FM.

For the psychological tests, it was found that patients endorsed significantly higher levels of depression, physical disability, sleep disturbance, poorer health status, and reduced activity levels in the exaggeration condition compared to the normal condition. These results provide partial support for the first hypothesis, which stated that patients would perform significantly worse on all measures when in the exaggeration condition. However, there was major overlap in scores obtained under the two conditions. No cutoff scores on the psychological tests could be identified that would accurately identify exaggeration.

Information has been extracted from numerous studies relating to the Beck Depression Inventory - II, FM Impact Questionnaire, Oswestry Disability Index - 2.0, and the Pittsburgh Sleep Quality Index for chronic pain patients and patients with FM (see Tables 33 to 36). These tables can be used to compare the scores obtained in this study to a number of known experimental and clinical groups. Scores on each of these measures from this study have been included in the tables to facilitate comparison across studies.

As seen in Table 33, patients in the normal condition of this study obtained scores similar on the Beck Depression Inventory - II to other studies involving patients with FM, and one study involving patients with Anxiety Disorders. Patients in the exaggeration condition obtained scores similar to studies involving university students faking depression. Notice, however, that there is considerable overlap in scores between subjects instructed to exaggerate or fake depression and actual patients with depressive disorders.
Table 33
Beck Depression Inventory-II scores from the literature

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Control Subjects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Students</td>
<td>120</td>
<td>12.6</td>
<td>9.9</td>
</tr>
<tr>
<td>Medical clinic (Outpatients)</td>
<td>340</td>
<td>8.7</td>
<td>9.7</td>
</tr>
<tr>
<td>Healthy Controls</td>
<td>21</td>
<td>7.0</td>
<td>9.1</td>
</tr>
<tr>
<td>Psychiatric Groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive Disorders</td>
<td>264</td>
<td>26.6</td>
<td>12.2</td>
</tr>
<tr>
<td>Anxiety Disorders</td>
<td>88</td>
<td>19.4</td>
<td>11.5</td>
</tr>
<tr>
<td>FM Patients:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM (Female)</td>
<td>64</td>
<td>19.7</td>
<td>11.9</td>
</tr>
<tr>
<td>FM (Outpatients)</td>
<td>23</td>
<td>20.0</td>
<td>10.1</td>
</tr>
<tr>
<td>Study 2 FM Patients (Normal condition)</td>
<td>49</td>
<td>20.0</td>
<td>10.8</td>
</tr>
<tr>
<td>Chronic Pain Patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rheumatoid Arthritis (Female)</td>
<td>30</td>
<td>13.6</td>
<td>7.9</td>
</tr>
<tr>
<td>Chronic Pain Disorders (Outpatients)</td>
<td>22</td>
<td>11.8</td>
<td>9.6</td>
</tr>
<tr>
<td>Analog Malingerers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fake Depression (University Students - Female)</td>
<td>26</td>
<td>33.8</td>
<td>13.8</td>
</tr>
<tr>
<td>Fake Depression (University Students – Male)</td>
<td>26</td>
<td>31.0</td>
<td>9.3</td>
</tr>
<tr>
<td>Study 2 FM Patients (Exaggeration Condition)</td>
<td>49</td>
<td>35.3</td>
<td>14.8</td>
</tr>
</tbody>
</table>

Note. All values are raw scores. ¹Beck et al. (1996), ²Arnau et al. (2001), ³Hassett et al. (2000), ⁴Suhr (2003), ⁵Lees-Haley (1989), ⁶Study 2 normal and exaggeration condition results.

As seen in Table 34, patients in the normal condition of this study obtained scores on the FM Impact Questionnaire that were similar to or higher than other patient groups reported in the literature. Patients in the exaggeration condition obtained scores similar to one sample of patients reported in the literature, but much higher than other patient groups. However, as illustrated in Table 34 and Figure 1, there is major overlap among the groups.
Table 34

Fibromyalgia Impact Questionnaire scores from the literature

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM Controls (Previously Undiagnosed with FM)</td>
<td>72</td>
<td>63.4</td>
<td>-</td>
</tr>
<tr>
<td>FM Patients (Previously Diagnosed with FM)</td>
<td>28</td>
<td>68.3</td>
<td>-</td>
</tr>
<tr>
<td>FM Patients (Placebo – Baseline)</td>
<td>23</td>
<td>52.7</td>
<td>14.3</td>
</tr>
<tr>
<td>FM Patients (Growth Hormone – Baseline)</td>
<td>22</td>
<td>50.0</td>
<td>13.1</td>
</tr>
<tr>
<td>FM Patients (Placebo – Finish)</td>
<td>23</td>
<td>46.6</td>
<td>17.8</td>
</tr>
<tr>
<td>FM Patients (Growth Hormone – Finish)</td>
<td>22</td>
<td>36.2</td>
<td>16.6</td>
</tr>
<tr>
<td>Study 2 FM Patients (Normal Condition)</td>
<td>49</td>
<td>60.4</td>
<td>16.2</td>
</tr>
<tr>
<td>Study 2 FM Patients (Exaggeration Condition)</td>
<td>49</td>
<td>73.2</td>
<td>15.4</td>
</tr>
</tbody>
</table>

Note. All values are raw scores. 1White, Nielson, Harth, Østbye, & Speechley (2002); Bennett et al. (1998): one group received growth hormone, one placebo injections, and scores were reported pre (baseline) and post (finish) injections. 2Study 2 normal and exaggeration condition results.

Figure 1

Fibromyalgia Impact Questionnaire scores in relation to the literature.

Note. The maximum score on the test is 80. The error lines represent two standard deviations. FM-1 = White et al., 2002, N = 72, no standard deviation was reported; FM-2 and FM-3 = Bennett et al., 1998, N = 23 and N = 22, baseline placebo and baseline growth hormone conditions; FM-Control and FM-Exaggeration are from the present study.

As seen in Table 35, the patients in the normal condition obtained scores on the Oswestry Disability Index - 2.0 similar to patients with FM who have low kinesiophobia (fear of movement), to chronic pain patients who are not working, and to chronic pain patients involved in the litigation process. Patients in the exaggeration condition obtained scores similar to studies involving patients with FM who have high kinesiophobia, and
Credibility and Fibromyalgia

chronic pain patients referred to a pain centre. Again, as illustrated in Table 35 and Figure 2, there is considerable overlap in the score distributions for the various clinical groups.

Table 35

Oswestry Disability Index 2.0 scores from the literature

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chronic Pain Patients:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM Patients (Female - Low Kinesiophobia)¹</td>
<td>143</td>
<td>40.2</td>
<td>12.9</td>
</tr>
<tr>
<td>FM Patients (Female - High Kinesiophobia)¹</td>
<td>90</td>
<td>47.3</td>
<td>13.5</td>
</tr>
<tr>
<td>Chronic Pain Patients (Pain Centre Referrals)²</td>
<td>424</td>
<td>47.6</td>
<td>-</td>
</tr>
<tr>
<td>Acute Low Back Pain (Outpatient)⁴</td>
<td>123</td>
<td>28.0</td>
<td>15.1</td>
</tr>
<tr>
<td>Chronic Low Back Pain (Outpatient)⁴</td>
<td>233</td>
<td>44.1</td>
<td>14.8</td>
</tr>
<tr>
<td>Study 2 FM Patients (Normal Condition)⁵</td>
<td>49</td>
<td>36.8</td>
<td>16.0</td>
</tr>
<tr>
<td>Study 2 FM Patients (Exaggeration Condition)⁵</td>
<td>49</td>
<td>56.4</td>
<td>17.8</td>
</tr>
<tr>
<td><strong>Employment Status:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic Pain (Not-Working)³</td>
<td>129</td>
<td>40.5</td>
<td>16.7</td>
</tr>
<tr>
<td>Chronic Pain (Working)³</td>
<td>117</td>
<td>25.9</td>
<td>14.6</td>
</tr>
<tr>
<td><strong>Litigation Status:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic Pain (Non-Litigation, Baseline)³</td>
<td>143</td>
<td>30.7</td>
<td>16.0</td>
</tr>
<tr>
<td>Chronic Pain (Litigation - In Progress, Baseline)³</td>
<td>103</td>
<td>40.1</td>
<td>15.7</td>
</tr>
<tr>
<td>Chronic Pain (Non-Litigation – Final Follow-up)³</td>
<td>50</td>
<td>30.7</td>
<td>18.8</td>
</tr>
<tr>
<td>Chronic Pain (Litigation-Settled, Final Follow-up)³</td>
<td>50</td>
<td>31.2</td>
<td>16.2</td>
</tr>
</tbody>
</table>

Note. All values are raw scores. Kinesiophobia is the fear of movement. ¹Turk, Robinson, & Burwinkle (2004), ²Wittink, Turk, Carr, Sukinnik, & Rogers (2004), ³Suter (2002), ⁴Grotle, Vøllestad, Veierød, & Brox (2004), ⁵Study 2 normal and exaggeration condition results.

Figure 2.

Oswestry Disability Index 2.0 scores in relation to the literature.

![Graph showing Oswestry Disability Index 2.0 scores](image)

Note. Scores range from 0-100%. The error lines represent two standard deviations. CP-Outpatients = Grotle et al., 2004, N = 233; CP-Not-Working, CP-Working, CP-No-Litigation Baseline, and CP-Litigation In Progress
Baseline = Suter, 2002, N = 129, N=117, N = 143, N=103; FM-Control and FM-Exaggeration are from the present study.

As seen in Table 36 and Figure 3, FM patients in both conditions obtained scores on the Pittsburgh Sleep Quality Index similar to other patient groups with sleep disorders or depression. Although patients in the exaggeration condition obtained the highest scores, there is major overlap amongst the groups.

Table 36
Pittsburgh Sleep Quality Index scores from the literature

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control Subjects:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy Subjects (Good Sleepers) (^1)</td>
<td>52</td>
<td>2.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Healthy Subjects (^2)</td>
<td>45</td>
<td>3.3</td>
<td>1.8</td>
</tr>
<tr>
<td><strong>Sleep Disordered Patients:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disorders of Initiating &amp; Maintaining Sleep Patients (^1)</td>
<td>45</td>
<td>10.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Disorders of Excessive Somnolence Patients (^1)</td>
<td>17</td>
<td>6.5</td>
<td>3.0</td>
</tr>
<tr>
<td>Primary Insomnia (Sleep Laboratory Referrals) (^2)</td>
<td>80</td>
<td>12.5</td>
<td>3.8</td>
</tr>
<tr>
<td><strong>Clinical Patients:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed Patients (Inpatients and Outpatients) (^1)</td>
<td>34</td>
<td>11.1</td>
<td>4.3</td>
</tr>
<tr>
<td>Study 2 FM Patients (Normal Condition) (^3)</td>
<td>49</td>
<td>11.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Study 2 FM Patients (Exaggeration Condition) (^3)</td>
<td>49</td>
<td>14.4</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Note. All values are raw scores. \(^1\)Buysse et al. (1989), \(^2\)Backhaus et al. (2002), \(^3\)Study 2 normal and exaggeration results.

Figure 3
Pittsburgh Sleep Quality Index scores from the literature.

Note. Scores range from 0-21. The error lines represent two standard deviations. Healthy Controls and Primary Insomnia = Backhaus et al., 2002, N = 45, N = 80; Depression = Buysse et al., 1989, N = 34; FM-Control and FM-Exaggeration are from the present study.
In the present study, a detailed examination of the Multidimensional Pain Inventory – 1 (MPI) scales revealed that patients performed significantly worse on four scales from Part 1 (pain severity, interference, life control, affective distress), one scale from Part 2 (negative responses), and one scale from Part 3 (household chores) in the exaggeration condition compared to the normal condition. The magnitude of difference between groups on Pain Severity was very large (Cohen's d=1.12), and the effect sizes for the other scales were medium to large.

As seen in Table 37, the score patterns in the control condition are similar to MPI score patterns reported in a previous study of patients with FM (Turk et al., 1996). However, the Turk et al. sample appeared to have higher scores on the “Support” scale, compared to the current patient sample. The “Support” scale is an appraisal of support received from a spouse, family, or significant other, and 19% of the current sample did not respond to this scale. Responses were missing for 14% of the Turk et al. sample for the MPI in general. The authors did not specify on which scales those missing responses occurred.

Information has been extracted from three studies relating to the MPI and patient groups with chronic pain or FM (see Tables 37 and 38). Table 37 includes one group of patients with FM, one group of chronic pain patients involved in the compensation process, and FM patients from Study 2. Table 38 includes a group of patients with FM, a subset of who have a comorbid diagnosis of depression, anxiety disorders, or no Axis I disorders. As seen in Tables 37 and 38, there is considerable overlap in MPI scores across groups. It would be very difficult to differentiate groups on the basis of these scores. There is some indication that FM patients instructed to exaggerate produced modestly worse scores on a few scales, such as pain severity and negative responses, than some of the other groups. However, the score with the biggest effect size in the present study (i.e., Pain Severity, d = 1.12) cannot accurately differentiate groups (as seen in Figure 4).
### Table 37
Multidimensional Pain Inventory – 1 test scores from the literature

<table>
<thead>
<tr>
<th>MPI Scale</th>
<th>FM (N=117)</th>
<th>Compensation (N=120)</th>
<th>Exaggeration (N=37)</th>
<th>Normal (N=37)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain Impact:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain Severity</td>
<td>4.1</td>
<td>3.9</td>
<td>4.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Interference</td>
<td>4.0</td>
<td>4.6</td>
<td>4.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Life Control</td>
<td>3.0</td>
<td>2.8</td>
<td>2.5</td>
<td>3.6</td>
</tr>
<tr>
<td>Affective Distress</td>
<td>3.5</td>
<td>3.6</td>
<td>4.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Support</td>
<td>4.1</td>
<td>4.6</td>
<td>3.5</td>
<td>3.4</td>
</tr>
<tr>
<td>Negative</td>
<td>2.0</td>
<td>1.9</td>
<td>2.8</td>
<td>2.1</td>
</tr>
<tr>
<td>Solicitous</td>
<td>3.0</td>
<td>3.6</td>
<td>2.7</td>
<td>2.9</td>
</tr>
<tr>
<td>Distracting</td>
<td>2.1</td>
<td>2.6</td>
<td>2.0</td>
<td>2.9</td>
</tr>
<tr>
<td>Household Chores</td>
<td>-</td>
<td>2.6</td>
<td>3.4</td>
<td>4.1</td>
</tr>
<tr>
<td>Outdoor Work</td>
<td>-</td>
<td>1.0</td>
<td>1.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Activities Away From</td>
<td>-</td>
<td>2.0</td>
<td>2.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Home</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Activities</td>
<td>-</td>
<td>2.4</td>
<td>2.3</td>
<td>2.6</td>
</tr>
<tr>
<td>General Activity Level</td>
<td>2.4</td>
<td>2.0</td>
<td>2.6</td>
<td>2.8</td>
</tr>
</tbody>
</table>

**Note.** All values are raw scores. Scores range from 0 to 6, with 6 representing the most severe response. 1Turk, Okifuji, Sinclair, & Starz (1996) – the missing scale scores were not reported, 2Rudy (1987), 3Study 2 exaggeration and normal condition results.

### Table 38
Multidimensional Pain Inventory – 1 test scores for patients with FM and comorbid depression and anxiety

<table>
<thead>
<tr>
<th>MPI Scale</th>
<th>Entire Sample (n=115)</th>
<th>Anxiety Disorder (n=33)</th>
<th>Mood Disorder (n=33)</th>
<th>No Axis 1 (n=46)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain Impact:</strong></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Pain Severity</td>
<td>4.1</td>
<td>1.0</td>
<td>4.5</td>
<td>.91</td>
</tr>
<tr>
<td>Interference</td>
<td>3.9</td>
<td>1.2</td>
<td>4.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Life Control</td>
<td>3.3</td>
<td>1.2</td>
<td>2.9</td>
<td>1.3</td>
</tr>
<tr>
<td>Affective Distress</td>
<td>3.2</td>
<td>1.3</td>
<td>3.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Support</td>
<td>3.7</td>
<td>1.7</td>
<td>4.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Negative</td>
<td>1.1</td>
<td>1.3</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Solicitous</td>
<td>3.1</td>
<td>1.7</td>
<td>3.6</td>
<td>1.4</td>
</tr>
<tr>
<td>Distracting</td>
<td>2.7</td>
<td>1.5</td>
<td>3.1</td>
<td>1.3</td>
</tr>
<tr>
<td>General Activity Level</td>
<td>3.7</td>
<td>1.4</td>
<td>3.1</td>
<td>1.4</td>
</tr>
</tbody>
</table>

**Note.** All values are raw scores. All scores are from Thieme, Turk, & Flor (2004). Four scales were not included in the report (i.e., household chores, outdoor work, activities away from home, and social activities).
Figure 4.
Pain Severity scores from the literature

![Graph showing pain severity scores from the literature]

Note. Scores range from 0-6. The error lines represent two standard deviations. FM = Turk et al., 1996, N = 117; FM Compensation = Rudy, 1987, N = 120; FM = Entire Sample, Thieme et al., 2004, N = 115; FM-Control and FM-Exaggeration are from the present study.

In the present study, patients obtained significantly higher scores on most of validity scales and indexes of the PAI (Inconsistency, Negative Impression Management, Defensiveness, Malingering, and Rogers scales), and a lower score on one scale (Positive Impression Management) in the exaggeration condition compared to the normal condition. There was no difference on the Infrequency scale. These results provide partial support for the first hypothesis. The Infrequency scale is used to identify individuals who respond in an atypical or random way, and patients scored in the normal range on this scale in both conditions. Therefore, the absence of a significant finding on this scale indicates that patients were not responding atypically to the PAI items in either experimental condition. The Inconsistency scale reflects the consistency with which the patient answered items with similar content. Although responses were elevated in the exaggeration condition, compared to the normal condition, mean scores were in the normal range for both conditions. Defensiveness, or the denial and minimization of psychopathology, is measured on the Positive Impression Management scale. Although scores were lower in the exaggeration condition for this scale, scores remained in the normal range for both conditions.

As for validity scales related to symptom exaggeration, the Negative Impression Management (NIM) scale consists of items that are associated with an unfavourable presentation, and are considered indicative of exaggerated psychopathology (Morey, 1991). Patients had significantly higher NIM scores in the exaggeration condition than in the normal
condition. The Malingering Index (MAL) and Rogers Discriminant Function (RDF) are also believed to measure symptom exaggeration (Liljequist et al., 1998; Rogers et al., 1996; Morey, 1996). Again, although scores on these scales and indexes were significantly higher in the exaggeration condition than in the normal condition, scores remained in the normal range for both conditions.

Patients obtained significantly higher scores on all of the clinical scales of the PAI in the exaggeration condition compared to the normal condition. However, those scores reached clinical significance (i.e., were beyond that reported by most people) on only some scales, according to the test manual (Morey, 1991). Mean scores were within the range indicating some distress on the Mania, Schizophrenia, and Borderline scales in the exaggeration condition. Elevated scores on the Borderline scale have been reported in a group of patients with work-related chronic musculoskeletal pain disability (Dersh et al., 2002), but elevated scores have not been reported previously on the Mania or Schizophrenia scales for related patient samples. Mean scores were in the range indicating some distress in both conditions on the Anxiety and Anxiety-Related Disorders scales. Mean scores were in the range indicating some distress in the normal condition, and they were in the range indicating prominent dysphoria in the exaggeration condition on the Depression scale. Finally, mean scores were in the range indicating significant distress in both conditions on the Somatic Complaints scale. The somatic, depressive, and anxiety-related scales score patterns in both the normal and exaggeration conditions are similar to findings in the medical and psychological literature that indicate a comorbidity of these issues and FM (Baumstark & Buckelew, 1992; Goldenberg, 1986; Hudson et al., 1992; Hudson et al., 1985; Slesinger et al., 2002; Smythe, 1989; Uveges et al., 1990; Walker et al., 1997a; Wolfe et al., 1990). The main difference is that patients in the exaggeration condition reported higher levels of distress on those scales.

Overall, results from the PAI suggest that when asked to exaggerate FM, patients obtained the highest scores on scales associated with their condition (i.e., the somatic complaints and depression scales). However, scores for all of the clinical scales increased significantly in the exaggeration condition, indicating exaggeration on a broad range of dimensions, not just dimensions related to FM. The fact that scores on the validity scales and indexes remained in the normal range for most patients (see Table 26), while clinical scale
scores were elevated in the exaggeration condition further suggests that patients with FM are able to exaggerate in a convincing manner.

PAI scores from this study can be compared to a number of groups from the literature in Tables 39 to 41. As seen in Table 39, scores on the Negative Impression Management (NIM) scale were mildly elevated for patients in the exaggeration condition, but roughly comparable to other clinical groups. Table 40 represents results from a study that involved coaching university students to feign that they were suffering from a mental disorder (Bagby et al., 2002). In the coached condition, participants were informed about the existence of the validity scales and given strategies on how to avoid detection, whereas participants in the uncoached condition were not informed of the validity scales. Scores in those conditions were compared with scores from patients with psychiatric disorders and a student control group. As seen in Tables 39-40, FM patients in the exaggeration condition of Study 2 scored a little lower on the NIM (but similarly) and much lower on the RDF compared to both the coached and uncoached malingerers on those scales. Lower scores on these validity scales suggest that FM patients in the exaggeration condition were better able to portray a convincing profile of disability than university students simulating mental disorder.
Table 39

Personality Assessment Inventory scores from the literature & Study 2

<table>
<thead>
<tr>
<th></th>
<th>Clinical&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Depressed&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Normal&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Exaggeration&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=1,246</td>
<td>N=126</td>
<td>N=49</td>
<td>N=49</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>SD</strong></td>
<td><strong>Mean</strong></td>
<td><strong>SD</strong></td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td><strong>Validity Scales &amp; Indexes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inconsistency</td>
<td>53 10</td>
<td>54 10</td>
<td>51.2 8.7</td>
<td>56.6 10.9</td>
</tr>
<tr>
<td>Infrequency</td>
<td>52 10.5</td>
<td>50 10.5</td>
<td>49.4 7.7</td>
<td>51.2 10.8</td>
</tr>
<tr>
<td>Negative Impression</td>
<td>62 14.5</td>
<td>65 14.5</td>
<td>58.4 11.4</td>
<td>67.1 17.4</td>
</tr>
<tr>
<td>Positive Impression</td>
<td>43 7.5</td>
<td>41 7.5</td>
<td>48.0 10.6</td>
<td>41.2 12.3</td>
</tr>
<tr>
<td>Defensiveness Index</td>
<td>-</td>
<td>-</td>
<td>51.5 10.3</td>
<td>47.4 13.0</td>
</tr>
<tr>
<td>Malingering Index</td>
<td>-</td>
<td>-</td>
<td>57.6 13.9</td>
<td>66.9 18.9</td>
</tr>
<tr>
<td>Rogers DF</td>
<td>-</td>
<td>-</td>
<td>49.2 9.5</td>
<td>53.5 10.7</td>
</tr>
<tr>
<td><strong>Clinical Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic Complaints</td>
<td>59 15</td>
<td>64 15</td>
<td>73.4 13.3</td>
<td>81.2 12.9</td>
</tr>
<tr>
<td>Anxiety</td>
<td>61 15</td>
<td>69 15</td>
<td>61.6 12.6</td>
<td>70.4 14.7</td>
</tr>
<tr>
<td>Anxiety Related Dis.</td>
<td>60 16</td>
<td>67 16</td>
<td>60.2 14.4</td>
<td>69.6 16.0</td>
</tr>
<tr>
<td>Depression</td>
<td>64 16</td>
<td>76 16</td>
<td>66.7 12.1</td>
<td>78.4 15.1</td>
</tr>
<tr>
<td>Mania</td>
<td>53 11</td>
<td>52 11</td>
<td>50.5 10.9</td>
<td>56.0 12.9</td>
</tr>
<tr>
<td>Paranoia</td>
<td>58 8.5</td>
<td>59 8.5</td>
<td>48.9 8.5</td>
<td>56.1 12.4</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>60 15</td>
<td>63 15</td>
<td>55.0 11.6</td>
<td>63.7 15.1</td>
</tr>
<tr>
<td>Borderline Features</td>
<td>64 15</td>
<td>68 15</td>
<td>54.0 11.5</td>
<td>62.4 12.6</td>
</tr>
<tr>
<td>Antisocial Features</td>
<td>56 13</td>
<td>54 13</td>
<td>46.1 7.9</td>
<td>51.2 10.3</td>
</tr>
<tr>
<td>Alcohol Problems</td>
<td>59 18</td>
<td>57 18</td>
<td>47.1 6.8</td>
<td>50.3 10.0</td>
</tr>
<tr>
<td>Drug Problems</td>
<td>58 17</td>
<td>55 12</td>
<td>51.1 9.8</td>
<td>56.2 12.1</td>
</tr>
</tbody>
</table>

Note. All values are T-scores. <sup>1</sup>Morey (1996), <sup>2</sup>Study 2 normal and exaggeration results. Rogers DF = Rogers Discriminant Function.

The results from a study in which students were asked to feign mental disorders are presented in Table 41. Scores on the Negative Impression Management scale (NIM) for FM patients in the exaggeration condition were comparable to sophisticated simulators feigning major depression and sophisticated simulators feigning mental disorders in general ("all disorders"), suggesting a similar response strategy in feigning depression between groups.

As for the FM patients in the normal condition, scores on the NIM scale were slightly lower than other clinical groups (see Table 39). In the study in which students were asked to feign mental disorders, NIM scores for FM patients in the normal condition were lower than the other groups feigning mental disorders, except for the generalized anxiety sophisticated simulators and control patient groups (Table 41), suggesting that patients with FM are unlikely to be falsely labeled as exaggerators.
Table 40

Personality Assessment Inventory scores from the literature

<table>
<thead>
<tr>
<th>Validity Scales &amp; Indexes</th>
<th>Coached N=22 Mean</th>
<th>SD</th>
<th>Uncoached N=22 Mean</th>
<th>SD</th>
<th>Patients N=75 Mean</th>
<th>SD</th>
<th>Student Controls N=45 Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative Impression</td>
<td>72.0</td>
<td>22.6</td>
<td>73.8</td>
<td>21.3</td>
<td>62.9</td>
<td>16.2</td>
<td>47.2</td>
<td>4.1</td>
</tr>
<tr>
<td>Rogers DF</td>
<td>72.2</td>
<td>12.4</td>
<td>67.9</td>
<td>10.6</td>
<td>50.6</td>
<td>9.1</td>
<td>51.4</td>
<td>7.0</td>
</tr>
<tr>
<td>Clinical Scales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic Complaints</td>
<td>66.7</td>
<td>17.6</td>
<td>70.9</td>
<td>15.7</td>
<td>69.3</td>
<td>15.0</td>
<td>46.0</td>
<td>5.3</td>
</tr>
<tr>
<td>Anxiety</td>
<td>69.9</td>
<td>16.0</td>
<td>74.3</td>
<td>16.8</td>
<td>67.6</td>
<td>16.7</td>
<td>50.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Anxiety Related Dis.</td>
<td>66.6</td>
<td>15.1</td>
<td>67.7</td>
<td>17.4</td>
<td>62.2</td>
<td>17.1</td>
<td>47.2</td>
<td>10.2</td>
</tr>
<tr>
<td>Depression</td>
<td>84.9</td>
<td>16.6</td>
<td>85.7</td>
<td>16.4</td>
<td>76.4</td>
<td>18.3</td>
<td>48.3</td>
<td>8.9</td>
</tr>
<tr>
<td>Mania</td>
<td>55.4</td>
<td>19.5</td>
<td>52.2</td>
<td>16.1</td>
<td>49.3</td>
<td>10.8</td>
<td>51.6</td>
<td>10.0</td>
</tr>
<tr>
<td>Paranoia</td>
<td>83.0</td>
<td>17.6</td>
<td>80.6</td>
<td>15.9</td>
<td>53.6</td>
<td>13.1</td>
<td>49.4</td>
<td>8.9</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>74.5</td>
<td>18.2</td>
<td>73.5</td>
<td>17.7</td>
<td>63.5</td>
<td>15.8</td>
<td>45.0</td>
<td>8.7</td>
</tr>
<tr>
<td>Borderline Features</td>
<td>69.8</td>
<td>10.1</td>
<td>72.7</td>
<td>11.8</td>
<td>61.1</td>
<td>14.2</td>
<td>53.0</td>
<td>10.1</td>
</tr>
<tr>
<td>Antisocial Features</td>
<td>61.4</td>
<td>19.0</td>
<td>60.0</td>
<td>15.7</td>
<td>49.4</td>
<td>10.1</td>
<td>52.9</td>
<td>9.7</td>
</tr>
<tr>
<td>Alcohol Problems</td>
<td>64.5</td>
<td>16.2</td>
<td>59.4</td>
<td>19.3</td>
<td>48.2</td>
<td>9.2</td>
<td>47.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Drug Problems</td>
<td>61.7</td>
<td>19.5</td>
<td>62.4</td>
<td>19.9</td>
<td>51.1</td>
<td>8.6</td>
<td>48.9</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Note. All values are T-scores. Abbreviations: Rogers DF = Rogers Discriminant Function. All scores are from Bagby et al. (2002): participants were either informed (coached) or not informed (uncoached) about the validity scales and strategies on how to avoid detection, and both groups were instructed to feign a mental disorder. These two groups were compared to psychiatric patients and an honest responding control group.

Table 41

Personality Assessment Inventory scores for Feigning Mental Disorders

<table>
<thead>
<tr>
<th>Simulators</th>
<th>Naive Mean</th>
<th>Sophisticated Mean</th>
<th>Controls</th>
<th>Patients Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Depression</td>
<td>n=39</td>
<td>n=21</td>
<td>n=136</td>
<td></td>
</tr>
<tr>
<td>Inconsistency</td>
<td>57.4</td>
<td>57.4</td>
<td>53.6</td>
<td></td>
</tr>
<tr>
<td>Negative Impression</td>
<td>86.0</td>
<td>68.6</td>
<td>64.9</td>
<td></td>
</tr>
<tr>
<td>Generalized Anxiety</td>
<td>n=38</td>
<td>n=19</td>
<td>n=40</td>
<td></td>
</tr>
<tr>
<td>Inconsistency</td>
<td>57.5</td>
<td>49.5</td>
<td>51.7</td>
<td></td>
</tr>
<tr>
<td>Negative Impression</td>
<td>75.7</td>
<td>54.8</td>
<td>53.5</td>
<td></td>
</tr>
<tr>
<td>All Disorders</td>
<td>n=39</td>
<td>n=21</td>
<td>n=136</td>
<td></td>
</tr>
<tr>
<td>Inconsistency</td>
<td>59.8</td>
<td>55.7</td>
<td>53.5</td>
<td></td>
</tr>
<tr>
<td>Negative Impression</td>
<td>83.9</td>
<td>68.8</td>
<td>63.5</td>
<td></td>
</tr>
</tbody>
</table>

Note. All values are T-scores. All scores are from Rogers et al. (1996).
As for the clinical scale scores, FM patients in the exaggeration condition scored considerably higher on the Somatic Complaints scale compared to the remaining groups in Tables 39 and 40. On the Depression scale the FM sample scored similarly to depressed patients (Table 39). Patients in the normal and exaggeration conditions scored lower on the Depression scale compared to the coached and uncoached groups feigning mental disorders (Table 40). As mentioned earlier, the response pattern on the clinical scales can provide some information regarding experimental condition, given that patients in the exaggeration condition scored higher on some of the clinical scales compared to other patient and subject groups. The validity scale scores of exaggerators were similar to other sophisticated groups instructed to feign mental disorders, and to depressed and anxiety patient groups, but slightly lower than naïve groups feigning mental disorders. The validity scale response patterns indicate that patients in the exaggeration will not likely be identified, based on performance of those validity scale scores.

The cutoff score analyses for the psychological tests and the Personality Assessment Inventory illustrates that performance in the exaggeration condition cannot be accurately differentiated from performance in the normal condition (see Table 26). A very limited number of patients exaggerating disability were correctly identified. These results do not support the second hypothesis, given that so few patients in the exaggeration condition could be correctly identified by the psychological and personality assessment cutoff scores, without grossly inflating the rate of incorrectly classifying those in the normal condition (false positives).

It has been suggested that multiple test score combinations may improve the identification rate of exaggeration over individual tests (Bianchini et al., 2005; Iverson et al., 1996; Meyers et al., 2002). Hence, a series of multiple cutoff score combinations were examined for some of the psychological tests and Personality Assessment Inventory (PAI) scales and indexes to determine whether they improved the effectiveness for identifying exaggeration over individual tests. The psychological test score combinations resulted in higher hit rates for exaggerators, compared to single test score cutoffs. However, this often came with the trade-off of higher false positives for non-exaggerators. Scoring above the cutoff on one or more of the three scale scores from the Multidimensional Pain Inventory (Pain Severity, Interference, and Affective Distress) resulted in a 73% hit rate for
exaggerators, the best hit rate of the psychological test combinations. Unfortunately, this hit rate resulted in 20% false positives (see Table 25). The best combination for scoring above the cutoff on two or more scores was the BDI, FIQ, ODI, and the MPI-Pain Interference scale combination, resulting in a 53% hit rate for exaggerators, with 4% false positives. Of course, nearly half of the exaggerators were not identified.

The PAI scale score combinations resulted higher hit rates than single test scores. Scoring above the cutoff on one or more of the combination involving the three validity scales (NIM, MAL, and RDF) and the selected clinical scales resulted in a 57% hit rate for exaggerators, the best hit rate of the PAI score combinations (see Table 31). However, this hit rate resulted in 22% false positives. This combination also resulted in the best combination for scoring above the cutoff on two or more scores on the PAI, with in a 45% hit rate for exaggerators and 10% false positives. Therefore, the psychological and PAI test combinations of two or more scores have clearly improved the hit rates for exaggerators, while keeping false positives below 10%. However, most exaggerators were not identified.

The classification rates from this study can be compared to a number of groups from the literature in Table 42. As seen in Table 42, the specificity rates obtained in the current study are similar to the majority of studies listed in the table, with the exception of the validity scale of the Pain Patient Profile, which was considerably lower than other studies. It is not unusual that the specificity rates are high among most of these studies, given that researchers typically have made it a priority to maximize specificity, thus minimizing false positives. The sensitivity of the Beck Depression Inventory - II and the Fibromyalgia Impact Questionnaire from the current study was lower than sensitivity obtained from most of the other studies, but similar to the McGill Pain Questionnaire, and the Pain Disability Index. The sensitivity obtained from the three combined MPI scale scores in the current study was similar to most of the studies in Table 42, and considerably higher than the sensitivity rates obtained by the McGill Pain Questionnaire, and the Meyers and Diep (2000) test combination. Even though the MPI scale combination had the best sensitivity and specificity in the current study, the sensitivity was still considerably lower than other studies (e.g., Meyers & Volbrecht, 2000; Meyers et al., 2002).
Table 42
Classification rates (%) from test scores of assessed malingering in pain patient samples from the literature

<table>
<thead>
<tr>
<th>Test Type</th>
<th>Cut Score</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGill Pain Questionnaire</td>
<td>49</td>
<td>21</td>
<td>90</td>
<td>10</td>
<td>79</td>
</tr>
<tr>
<td>Pain Disability Index</td>
<td>54</td>
<td>59</td>
<td>90</td>
<td>10</td>
<td>41</td>
</tr>
<tr>
<td>Modified Somatic Perception Questionnaire 1</td>
<td>10</td>
<td>90</td>
<td>90</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Depression Scale - Pain Patient Profile</td>
<td>60</td>
<td>77</td>
<td>77</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>Anxiety Scale - Pain Patient Profile</td>
<td>58</td>
<td>69</td>
<td>73</td>
<td>23</td>
<td>31</td>
</tr>
<tr>
<td>Somatization Scale - Pain Patient Profile 2</td>
<td>57</td>
<td>85</td>
<td>87</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Validity Scale - Pain Patient Profile</td>
<td>10</td>
<td>69</td>
<td>81</td>
<td>19</td>
<td>31</td>
</tr>
<tr>
<td>Depression Scale - Pain Patient Profile</td>
<td>58</td>
<td>70</td>
<td>85</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Anxiety Scale - Pain Patient Profile</td>
<td>56</td>
<td>65</td>
<td>72</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>Somatization Scale - Pain Patient Profile 3</td>
<td>56</td>
<td>60</td>
<td>87</td>
<td>13</td>
<td>40</td>
</tr>
<tr>
<td>Validity Scale - Pain Patient Profile</td>
<td>8</td>
<td>65</td>
<td>20</td>
<td>80</td>
<td>35</td>
</tr>
<tr>
<td>Fail 2 or more Neuropsychological Tests 4</td>
<td>-</td>
<td>29</td>
<td>100</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>Fail 2 or more Neuropsychological Tests 5</td>
<td>-</td>
<td>83</td>
<td>100</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Combined Validity Index (MMPI-2) 6</td>
<td>5</td>
<td>86</td>
<td>100</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Beck Depression Inventory -II – Study 2 7</td>
<td>36</td>
<td>43</td>
<td>92</td>
<td>8</td>
<td>57</td>
</tr>
<tr>
<td>Fibromyalgia Impact Questionnaire – Study 2 7</td>
<td>78.9</td>
<td>41</td>
<td>94</td>
<td>6</td>
<td>59</td>
</tr>
<tr>
<td>PAI: NIM scale – Study 2 7</td>
<td>77</td>
<td>20</td>
<td>94</td>
<td>6</td>
<td>80</td>
</tr>
<tr>
<td>Failed 1 or more (PS, I, &amp; AD) – Study 2 7</td>
<td>-</td>
<td>73</td>
<td>80</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>Failed 1 or more (PAI: NIM, MAL, &amp; 6 clinical scales) – Study 2 7</td>
<td>-</td>
<td>57</td>
<td>78</td>
<td>22</td>
<td>43</td>
</tr>
</tbody>
</table>

Note. Abbreviations: FP = False Positives, FN = False Negatives.
1. Larabee (2003) - litigants meeting the criteria for either definite or probable malingered neurocognitive dysfunction and chronic pain were compared across three self-report pain measures.
2. McGuire et al. (2001) – a group of pain patients (pain controls) were compared to chronic pain patients instructed to exaggerate their pain (pain simulators) on the Pain Patient Profile.
3. McGuire & Shores (2001) – a group of pain patients (pain controls) v. pain simulators (students instructed to feign a pain disorder as part of a compensation case) were compared on the Pain Patient Profile.
5. Meyers & Volbrecht (2003) – examined whether a combination of 9 neuropsychological tests could correctly identify litigant and nonlitigant groups consisting of traumatic brain-injured patients, some of whom also had chronic pain or depression, community controls, and students feigning a brain injury in a compensation situation.
7. Study 2 normal and exaggeration results using scores from the Beck Depression Inventory, the Fibromyalgia Impact Questionnaire, the Multidimensional Pain Inventory scale combination (PS = Pain Severity, I = Interference, AD = Affective Distress), and a Personality Assessment Inventory scale and combination (Negative Impression Management, Malingering Index, Depression, Anxiety, Anxiety Related, Somatic Complaints, Schizophrenia, and Borderline scales).

Patients did not perform worse in the exaggeration condition on the three effort tests (Victoria Symptom Validity Test (VSVT), Test of Memory Malingering (TOMM), and 21 Item Test), compared to the normal condition. In fact, most patients scored well above the
cutoff on these measures, indicating adequate effort. These results do not support the first hypothesis. Of the total sample, only 12 patients indicated that they had exaggerated on the effort tests. After eliminating those patients who indicated that they had not exaggerated on the effort tests from the analyses, 58% of patients in the exaggeration condition scored below the cutoffs on measures of the VSVT, compared with 31% of the total patient sample. As for measures on the TOMM, 33% of those patients who indicated that they exaggerated scored below the cutoffs, compared with 11% of the total patient sample. There was virtually no difference between conditions for exaggerators on the 21 Item Test scores (Table 20). It is not surprising that a greater percentage of patients in the sample that exaggerated on the effort tests scored below the cutoffs on two of those tests, compared to the total patient sample. However, a large percentage of patients who said they exaggerated went undetected on the effort tests.

The effort test scores from this study can be compared to a number of groups from the literature in Tables 43 to 45. As seen in Table 43, all patients and exaggerators in the normal and exaggeration conditions obtained Easy Correct scores on the Victoria Symptom Validity Test similar to all groups in the table, except for experimental-malingerers. All patients and exaggerators in the normal condition, as well as all patients in the exaggeration condition, obtained Hard Correct and Total Correct scores that were similar to all groups except for compensation-seeking patients and experimental-malingerers. Exaggerators in the exaggeration condition obtained Hard Correct and Total Correct scores that were similar to compensation-seeking patients and experimental-malingerers. In the present study, most patients instructed to exaggerate did not provide poor effort on the VSVT.
Table 43

Victoria Symptom Validity Test scores from the literature

<table>
<thead>
<tr>
<th>Group</th>
<th>Group</th>
<th>N</th>
<th>Easy Correct</th>
<th>Hard Correct</th>
<th>Total Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>1</td>
<td>95</td>
<td>24.0 (.2)</td>
<td>23.4 (.9)</td>
<td>47.4 (.9)</td>
</tr>
<tr>
<td>Controls</td>
<td>2</td>
<td>21</td>
<td>24.0 (0)</td>
<td>23.3 (.9)</td>
<td>na</td>
</tr>
<tr>
<td>Clinical Patients (Non Forensic)</td>
<td>3</td>
<td>32</td>
<td>23.5 (1.2)</td>
<td>22.6 (1.8)</td>
<td>46.2 (2.6)</td>
</tr>
<tr>
<td>Clinical Patients (Non Forensic)</td>
<td>4</td>
<td>27</td>
<td>24.0 (0)</td>
<td>23.4 (1.05)</td>
<td>na</td>
</tr>
<tr>
<td>Moderate-Severe TBI Patients</td>
<td>5</td>
<td>30</td>
<td>23.8 (.7)</td>
<td>22.5 (2.6)</td>
<td>na</td>
</tr>
<tr>
<td>Clinical Patients (Non Forensic)</td>
<td>6</td>
<td>30</td>
<td>23.9 (.3)</td>
<td>23.2 (1.4)</td>
<td>na</td>
</tr>
<tr>
<td>Compensation-Seeking Patients</td>
<td>7</td>
<td>151</td>
<td>23.3 (2.0)</td>
<td>21.3 (3.9)</td>
<td>44.6 (5.4)</td>
</tr>
<tr>
<td>Compensation-Seeking Patients</td>
<td>8</td>
<td>205</td>
<td>23.3 (2.0)</td>
<td>20.2 (4.8)</td>
<td>43.5 (6.1)</td>
</tr>
<tr>
<td>Compensation-Seeking Patients</td>
<td>9</td>
<td>53</td>
<td>22.7 (2.5)</td>
<td>17.0 (5.8)</td>
<td>na</td>
</tr>
<tr>
<td>Experimental-Malingers</td>
<td>10</td>
<td>21</td>
<td>22.8 (2.2)</td>
<td>16.8 (6.8)</td>
<td>39.6 (8.3)</td>
</tr>
<tr>
<td>Experimental-Malingers</td>
<td>11</td>
<td>43</td>
<td>20.3 (4.4)</td>
<td>11.0 (6.1)</td>
<td>31.3 (9.1)</td>
</tr>
<tr>
<td>Experimental-Malingers</td>
<td>12</td>
<td>20</td>
<td>20.9 (3.5)</td>
<td>10.9 (6.7)</td>
<td>na</td>
</tr>
<tr>
<td>Study 2 FM Patients (Normal Condition)</td>
<td>13</td>
<td>49</td>
<td>23.7 (.8)</td>
<td>22.5 (2.3)</td>
<td>46.2 (2.8)</td>
</tr>
<tr>
<td>Study 2 FM Patients (Exaggeration Condition)</td>
<td>14</td>
<td>49</td>
<td>23.3 (2.1)</td>
<td>20.3 (5.5)</td>
<td>43.6 (7.1)</td>
</tr>
<tr>
<td>Study 2 FM Patients-E (Normal Condition)</td>
<td>15</td>
<td>12</td>
<td>23.9 (.29)</td>
<td>23.0 (1.2)</td>
<td>46.9 (1.3)</td>
</tr>
<tr>
<td>Study 2 FM Patients-E (Exaggeration Condition)</td>
<td>16</td>
<td>12</td>
<td>23.0 (1.9)</td>
<td>15.7 (6.5)</td>
<td>38.7 (7.7)</td>
</tr>
</tbody>
</table>

Note. The average number correct is reported first and the standard deviation is reported in parentheses.
Abbreviations: TBI = Traumatic Brain Injury. 1Slick et al. (1997), 2Strauss et al. (1999), 3Tan et al. (2002),
4Grote et al. (2000), 5Doss, Chelune, & Naugle (1999), 6Study 2 normal and exaggeration results. E=sample of
patients who exaggerated on effort tests (n=12).

Table 44 reveals that all patients and exaggerators in the normal and exaggeration
conditions obtained scores on the Test of Memory Malingering similar to control and patients
groups, except for elderly patients with dementia. Test scores were higher in both conditions
of the current study, compared to all subject groups instructed to malinger. These results are
interesting because many clinicians believe that patients who are in pain, are depressed, or
both might fail effort tests due to distraction and poor concentration. In the present study, not
a single patient failed the TOMM in the normal condition.
### Table 44

Test of Memory Malingering scores from the literature

<table>
<thead>
<tr>
<th>Group</th>
<th>Source</th>
<th>N</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy Control Subjects</td>
<td>Tombaugh (1997)</td>
<td>70</td>
<td>47.8</td>
<td>49.9</td>
<td>49.9</td>
</tr>
<tr>
<td>Community Controls</td>
<td>Tombaugh (1997)</td>
<td>70</td>
<td>(2.4)</td>
<td>(0.4)</td>
<td>(0.5)</td>
</tr>
<tr>
<td>Community Controls</td>
<td>Bolan et al. (2002)</td>
<td>16</td>
<td>n/a</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>University Students</td>
<td>Tan et al. (2002)</td>
<td>27</td>
<td>48.3</td>
<td>49.6</td>
<td>49.6</td>
</tr>
<tr>
<td>University Students</td>
<td>Powell et al. (2004)</td>
<td>28</td>
<td>48.8</td>
<td>49.9</td>
<td>49.9</td>
</tr>
<tr>
<td>University Students (Feedback)</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>n/a</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>University Students (No Feedback)</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>n/a</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>University Students (Computerized Version)</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>49.9</td>
<td>50.0</td>
<td>(0.0)</td>
</tr>
<tr>
<td>Elderly Community Subjects</td>
<td>Ashendorf et al. (2004)</td>
<td>197</td>
<td>48.9</td>
<td>49.9</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Patient Samples**

<table>
<thead>
<tr>
<th>Mixed Clinical Sample (Cognitively Impaired)</th>
<th>Tombaugh (1997)</th>
<th>42</th>
<th>43.9</th>
<th>48.6</th>
<th>49.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with Aphasia</td>
<td>Tombaugh (1997)</td>
<td>21</td>
<td>46.3</td>
<td>49.3</td>
<td>49.8</td>
</tr>
<tr>
<td>Patients with Traumatic Brain Injuries (TBI)</td>
<td>Tombaugh (1997)</td>
<td>45</td>
<td>45.9</td>
<td>49.4</td>
<td>49.6</td>
</tr>
<tr>
<td>Patients with Dementia</td>
<td>Tombaugh (1997)</td>
<td>37</td>
<td>41.0</td>
<td>45.7</td>
<td>47.0</td>
</tr>
<tr>
<td>Community Subjects (TBI)</td>
<td>Rees et al. (1998)</td>
<td>10</td>
<td>47.5</td>
<td>49.6</td>
<td>49.6</td>
</tr>
<tr>
<td>Patients (TBI)</td>
<td>Rees et al. (1998)</td>
<td>13</td>
<td>45.3</td>
<td>49.5</td>
<td>50.0</td>
</tr>
<tr>
<td>Mixed Sample of Hospital Inpatients</td>
<td>Rees et al. (1998)</td>
<td>13</td>
<td>47.9</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>Elderly Patients (Dementia)</td>
<td>Teichner &amp; Wagner (2004)</td>
<td>21</td>
<td>36.9</td>
<td>39.5</td>
<td>39.5</td>
</tr>
<tr>
<td>Elderly Patients (Cognitive Impairment)</td>
<td>Teichner &amp; Wagner (2004)</td>
<td>36</td>
<td>44.4</td>
<td>48.6</td>
<td>48.3</td>
</tr>
<tr>
<td>Elderly Patients (No Cognitive Impairment)</td>
<td>Teichner &amp; Wagner (2004)</td>
<td>21</td>
<td>47.4</td>
<td>49.7</td>
<td>49.7</td>
</tr>
<tr>
<td>Patients (Temporal Lobe Epilepsy)</td>
<td>Hill et al. (2003)</td>
<td>48</td>
<td>45.7</td>
<td>49.3</td>
<td>49.0</td>
</tr>
<tr>
<td>Psychiatric Inpatients (Major Depression)</td>
<td>Rees et al. (2001)</td>
<td>26</td>
<td>47.8</td>
<td>49.9</td>
<td>49.9</td>
</tr>
<tr>
<td>Study 2 FM Patients (Normal Condition)</td>
<td>Current Study</td>
<td>54</td>
<td>48.8</td>
<td>49.8</td>
<td>49.6</td>
</tr>
<tr>
<td>Study 2 FM Patients -E (Normal Condition)</td>
<td>Current Study</td>
<td>12</td>
<td>48.7</td>
<td>50.0</td>
<td>49.7</td>
</tr>
</tbody>
</table>

**Subjects Instructed to Malinger**

<table>
<thead>
<tr>
<th>University Students</th>
<th>Tombaugh (1997)</th>
<th>20</th>
<th>27.2</th>
<th>27.9</th>
<th>26.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>University Students</td>
<td>Tan et al. (2002)</td>
<td>25</td>
<td>34.5</td>
<td>36.5</td>
<td>35.3</td>
</tr>
<tr>
<td>University Students (Symptom Coached)</td>
<td>Powell et al. (2004)</td>
<td>27</td>
<td>26.6</td>
<td>26.7</td>
<td>26.7</td>
</tr>
<tr>
<td>University Students (Test Coached)</td>
<td>Powell et al. (2004)</td>
<td>25</td>
<td>32.3</td>
<td>34.6</td>
<td>34.5</td>
</tr>
<tr>
<td>Community Subjects</td>
<td>Bolan et al. (2002)</td>
<td>16</td>
<td>n/a</td>
<td>27.4</td>
<td>27.6</td>
</tr>
</tbody>
</table>
Credibility and Fibromyalgia

<table>
<thead>
<tr>
<th>Group</th>
<th>Source</th>
<th>N</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>University Students (Feedback)(^1)</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>n/a</td>
<td>29.9</td>
<td>28.3</td>
</tr>
<tr>
<td>University Students (No Feedback)(^1)</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>n/a</td>
<td>27.9</td>
<td>26.7</td>
</tr>
<tr>
<td>University Students (Computerized Version)</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>n/a</td>
<td>25.1</td>
<td>23.6</td>
</tr>
<tr>
<td>University Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects with TBIs</td>
<td>Bolan et al. (2002)</td>
<td>10</td>
<td>(3.6)</td>
<td>(4.1)</td>
<td>(6.1)</td>
</tr>
<tr>
<td>Study 2 FM Patients (Exaggeration)</td>
<td>Rees et al. (1998)</td>
<td>25</td>
<td>32.6</td>
<td>35.7</td>
<td>34.3</td>
</tr>
<tr>
<td>Study 2 FM Patients -E (Exaggeration)</td>
<td>Current Study</td>
<td>8</td>
<td>28.1</td>
<td>32.1</td>
<td>31.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
<td>47.2</td>
<td>48.6</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>46.2</td>
<td>46.7</td>
<td>46.7</td>
</tr>
</tbody>
</table>

Note. E=sample of patients who exaggerated on effort tests (n=12). Most of this table was derived from Iverson, Le Page, Koehler, Shojania, and Badii (under review).

\(^1\)The university students were instructed to fake memory problems. While taking the test, some were given feedback following each item as to whether it was correct or incorrect, and others received no feedback.

\(^2\)The university students were given a week to prepare, using any materials they wanted, to malinger believable problems relating to a head injury sustained in a motor vehicle accident.

\(^3\)Symptom coached students were told to fake cognitive impairment and they were given information describing typical symptoms and problems in people with traumatic brain injuries. Test coached students were given instructions describing effective test-taking strategies that would help them avoid detection.

As seen in Table 45, all patients and exaggerators in the normal and exaggeration conditions obtained scores on the 21 Item Test similar to all groups in the table, except for experimental, student, and psychiatric malingerer samples. The patients with FM were not falsely labeled as giving poor effort in the normal condition. As with the other effort tests, nearly all of the patients with FM instructed to exaggerate did not provide poor effort on this test.
Overall, results from the effort tests suggest that when asked to exaggerate, FM patients did not obtain scores similar to subject groups instructed to mangle or groups that typically obtain lower scores (i.e., compensation-seeking groups). Effort test scores remained in the normal range, while psychological test scores were elevated in the exaggeration condition, suggesting that patients with FM are able to exaggerate in a convincing manner.

Failure rates from the entire patient sample on the effort tests were similar in the exaggeration and control conditions. These results are contrary to a previous study that reported a significant difference in effort test scores between a group of FM patients involved in compensation or disability claims, and a comparison group of FM patients not involved in disability claims (Gervais et al., 2001). Gervais and colleagues noted that a greater number of
patients involved in a disability claim (35%) had failed the effort tests, compared to the failure rate of patients not involved in a disability claim (4%). The current results are unusual, given that the vast majority of FM patients did not fail the effort tests even when they were instructed to exaggerate.

Gervais et al. (2001) indicated that the test administrator was not blind to the experimental hypothesis or patient group membership, suggesting a possible biasing or experimenter effect that may have influenced performance on the effort tests. The research assistants administering the tests to patients in the current study were blind to the experimental condition, thereby avoiding the possibility of experimenter effects.

The current results are similar to the findings described earlier of McGuire and Shores (2001) who found poor identification rates between pain simulators and pain patients on the validity scale of the Pain Patient Profile. The validity scale had a moderate sensitivity rate (65%), but an extremely low specificity rate (20%). A related study using the same measure reported a similar sensitivity rate (69%) and an improved specificity rate (81%) on the validity scale, between a group of pain patients instructed to exaggerate their pain and a pain control group (McGuire et al., 2001). However, McGuire et al. found that the sensitivity (53%) and specificity (58%) rates were no better than chance on the validity scale, when comparing the pain patients instructed to exaggerate pain with a group of non-pain patients simulating pain. Fishbain et al. (2002) compared one group of chronic pain patients with secondary gain factors and another group with no secondary gain factors, pre or post treatment. They found that the secondary gain group did not differentially respond on an illness exaggeration scale compared to the control group. The findings from these studies are similar to the current study because clinical scales were found to be more successful at differentiating exaggeration in chronic pain patients than validity scales or effort tests. However, the clinical scales were only moderately effective.

One study has examined a combination of seven different validity scales of the MMPI-2 in assessing exaggeration in chronic pain patients (Meyers et al., 2002). The combination of validity scales results in a common weighted score, and this method was able to correctly identify 100% of nonlitigant chronic pain patients and 86% of a group instructed to exaggerate pain for compensation purposes. The results from this study show the most promise among validity scales in the identification of exaggeration in chronic pain.
populations. In contrast, the best test score combination from the current study resulted in a correct identification rate of 80% for nonexaggerators and 73% for exaggerators.

In the post-experimental interview, the majority of patients (91%) indicated that they tried to exaggerate disability overall during the exaggeration condition, while 74% reported at least some difficulties exaggerating. Patients identified several reasons behind those difficulties, including not feeling comfortable or feeling it was wrong to exaggerate, and not being familiar with exaggerating their symptoms (noting that they typically try to ignore or understate their symptoms). Of note, several patients in the study spontaneously reported difficulty exaggerating after the exaggeration condition. They elaborated by indicating that exaggerating the FM could create further difficulties for patients with FM, given the disbelief, confusion, and lack of understanding that already surrounds the condition. Patients further suggested that they frequently feel invalidated by others regarding their pain and disability. In fact, two patients spontaneously provided the experimenter with a written documentary of their experiences with FM over the years. The spontaneous patient reports are similar to the comments of Crombez et al. (2004), who suggested that FM patients may consider the origin of their pain as more mysterious than chronic low back pain. Patients with FM may experience a greater sense of discomfort in exaggerating something that they believe has a mysterious and invalidating quality attached to it.

A small minority of patients (4%) indicated that they did not exaggerate on the psychological questionnaires, while the majority of patients (65%) indicated that they either did not exaggerate, or did not know how to exaggerate, on the effort tests. The majority of patients (79%) rated their ability to exaggerate disability on the psychological questionnaires as a 5 or greater (on a scale of 1 to 10), while the majority of patients (80%) gave themselves a rating below 5 on the effort tests. These ratings are reflective of the overall results, which revealed that patients did exaggerate on the psychological and personality questionnaires, but not on the effort tests.

Patient explanations for the lack of exaggeration on effort tests included: (1) forgetting to exaggerate, (2) not knowing how to exaggerate, (3) not needing to exaggerate because the effort tests were already difficult, (4) not experiencing memory problems as part of FM, and (5) an inability to exaggerate when someone else was present. The first and second patient explanations are not supported by previous research, which has tended to
demonstrate that participants are able to exaggerate with little or no difficulty. However, instructions incorporated in previous research have tended to specify the exaggeration of cognitive impairments (e.g., Iverson et al., 1991; Iverson et al., 1994; Iverson & Franzen, 1996; Iverson, 1998; Rees et al., 1998; Tombaugh, 1997; Thompson, 2002), whereas the current study involved instructions to exaggerate problems associated with FM. The third patient explanation is not supported by the psychological research, which indicates that patients with FM report memory problems (e.g., Suhr, 2003). It has also been noted that patients may fail one type of effort test more than another because of the relevance of the test material to their presenting complaints (Tombaugh, 1996). Meyers and Volbrecht (2003) concurred and suggested the importance of using techniques specific to the nature of the alleged dysfunction in the detection of exaggeration and malingering, rather than one global measure. The current sample of FM patients was not specifically instructed to exaggerate cognitive impairments. Hence, some FM patients may not have exaggerated on the effort tests because they were not specifically instructed to do so.

The fourth stated reason for the lack of exaggeration on effort tests was an inability to exaggerate when someone else was present. Although only a small percentage of patients acknowledged the fourth reason as an explanation, this type of subject feedback has been raised in previous research. In a study examining a number of effort tests, some student participants reported difficulty maintaining their strategy to malinger and found it hard to give wrong answers (Tan, Slick, Strauss, & Hultsch, 2002). Two malingerers spontaneously reported difficulty giving wrong responses when tests were administered by the experimenter, compared to computerized tests. The paper and pencil version of the TOMM (experimenter administered) was given, while computerized versions of the VSVT and the Word Memory Test were given by Tan et al. Even though the experimenter was present during all of the testing, they were face-to-face with subjects for the administration of the TOMM, whereas they were physically present but sat at a distance and did not interact with subjects for the remaining effort tests.

In the current study, a research assistant was present and interacting with the patient during the effort tests, while patients worked independently for the remainder of the tests. The TOMM was also administered by a research assistant in the current study, but unlike Tan et al. (2002), a research assistant sat next to the patient during the computerized
administration of the VSVT, and read the test instruction from the monitor to the patient (as instructed in the VSVT manual). In addition, the 21 Item Test was administered by a research assistant in a face-to-face manner in the current study, compared to the computerized administration of the Word Memory Test in the Tan et al. study. Overall, there was more direct interaction and greater contact with the research assistant during administration of the effort tests in the current study. It may be that a greater number of patients in the current study experienced difficulties exaggerating when someone else was present, but they may not have acknowledged, or been aware of such effects.

The literature related to the patient perspective and chronic pain tends to suggest that patients struggle with issues of credibility and the invalidation of their pain and suffering, and the perception by others that they may be “malingersers” (Glenton, 2003; Richardson, 2005). It is possible that the subjects in the current study were sensitive to those perceptions, and that may have made them fundamentally uncomfortable exaggerating. In fact, when patients were asked in the post-experimental interview what difficulties they had overall exaggerating, they indicated that they felt uncomfortable or felt it was wrong to exaggerate.

According to the literature, a substantial minority of people involved in personal injury litigation or disability claims fail effort tests. These results have been found with a number of patient groups, including people with mild head injuries (Binder, 1993), chronic fatigue syndrome (van der Werf et al., 2000), chronic pain following whiplash injuries (Schmand et al., 1998), and in patients with FM (Gervais et al., 2001). These studies have incorporated groups of subjects involved in the compensation-seeking process, whereas this was not the case for the current patient sample. The current patient sample was obtained from two private rheumatology practices and one outpatient hospital program, and did not include a group of patients specifically involved in litigation. In total, a life, disability, health, or worker’s compensation claim was in progress for 26% of the total patient sample during the time of the study, and FM was included in only 4% of those claims. FM was not part of the claim for the remaining 26% of patients involved in an insurance claim. Hence, there was a virtual absence of FM-related compensation-seeking patients in the current sample. Contrary to the studies that included a homogenous group of patients seeking compensation for a specific medical condition, the majority of compensation-seeking patients in the current sample were not homogenous in that respect. For example, one patient was seeking
compensation for a wrist injury. In an attempt to simulate the incentives encountered in the claims process, researchers have instructed subjects to pretend that they are part of a compensation case (e.g., McGuire et al., 2001; McGuire & Shores, 2001). An association with claims process was not incorporated into the experimental instructions in the current study. Therefore, if involvement in the compensation-seeking process, or instructions to simulate such influences plays a role in the performance on effort tests, then the near absence of these influences in the current study may have contributed to the lack of an exaggeration effect on the effort tests.

In terms of the Insurance Inventory and issues of exaggeration, patients were asked what would happen to health policy coverage if they exaggerated their medical symptoms. The majority of patients (76%) indicated a negative outcome, resulting in the policy being cancelled, denied, or in a penalty being incurred. A minority of patients indicated that either nothing would happen (7%), or that more coverage would be given (2%). These results suggest that the majority of the patient sample have knowledge regarding the effects that exaggeration could have on insurance policy coverage.

**Limitations and Considerations for Future Research**

Research participants recruited from rheumatology clinics might have more severe problems than patients seen in other clinical and community settings (Wolfe et al., 1997a). Therefore, the current patient sample may not be representative of all patients with FM because they were recruited from private rheumatology clinics, which may be a limiting factor in the generalizability of the findings to other settings. Another limitation related to the patient sample is the small number of men included in the study, which limits generalizability. This is a common problem, of course, because women greatly outnumber men with this condition. The generalizability of these results is also limited by the laboratory/analog design of the study and how subjects might differ in a "real world" setting.

A consideration for a future study would be the comparison of effort test results that are experimenter-administered compared to tests that are self-administered. A comparison of test administration techniques would help clarify the issue of whether patients respond differently, depending on level of experimenter involvement.

In terms of subject sample, a direction for future research would to study various types of litigants, for example, the inclusion of individuals in the litigation process for health-
related claims, property-related claims, and employment-related claims. This could assist in determining whether the process of litigation and the associated distress is related to performance on effort tests, or if performance on effort tests is related to involvement in only specific areas of litigation and the compensation process. A third consideration related to the compensation-seeking process would be a modification of the current experimental instructions. As mentioned earlier, instructions to simulate such influences may affect a patient's performance on effort tests, and those influences were absent in the current experimental instructions. Therefore, instructing patients to exaggerate their medical condition in anticipation of a compensation claim would allow a further examination of the role that the compensation-seeking process plays in effort test performance for patients with FM.

Another consideration related to subject sample would be the inclusion of a group of patients who were diagnosed with FM by their GP, and who have never been treated by a rheumatologist. As suggested by Wolfe et al. (1997a), the inclusion of such a group of patients with FM would broaden the range of FM severity in the sample, and hence be more representative of patients with FM.
Credibility and Fibromyalgia

**General Discussion**

Bennett (1996) has summarized the main problems encountered in assessing the patient with chronic pain as follows: (a) pain is a subjective experience, (b) chronic pain cannot be completely understood in terms of classical medical models of disease that equate pathogenesis with physical pathology or dysfunction, (c) many people with chronic pain are not disabled, and (d) disability due to pain results from a variety of factors, including past experiences, motivation, personal value systems, cultural background, education, and psychological distress. Consequently, chronic pain conditions like FM present a special challenge to the medical system, the insurance industry, and the courts.

The research on FM has generally maintained a narrow focus on individual theoretical models, with no attempt at integrating those models. Numerous investigations have been conducted from a variety of perspectives in the medical and psychological research, but all have been inconclusive thus far. FM is one of the most problematic conditions for the insurance industry and the courts, due mainly to the substantial impact it can have on an individual’s ability to work. The courts have not reached a consensus regarding the validity of FM, and as such, plaintiff credibility is an important issue. The purpose of this dissertation was to examine the complexities faced by patients with FM, the medical profession, the insurance industry, and the legal system. A multidisciplinary approach has been taken to appreciate the diverse perspectives that are part of disability determinations in patients with FM, and to examine how each perspective may interface with the other. Two independent, yet related studies comprised this dissertation, each examining different aspects of disability claims in patients with FM.

Study 1 was an exploratory case review study of litigated cases involving FM in the Canadian courts. A number of variables were examined in the case judgments, including factors related to demographic, insurance, and credibility issues. The study revealed that the gender disparity found in the medical literature also appears to be present within the legal context. The study also revealed that although the presence of investigative and surveillance information alone did not have an impact on the outcome of disability determinations for FM, the credibility of that information was central to the importance the judge placed on that information. “Persuasive” or credible surveillance information resulted in awards that were
20-30% the size of the awards granted to such cases in which the surveillance information was not considered credible.

Of course, credibility played a central role in disability determinations for the plaintiff, and significantly affected the outcome. Plaintiffs perceived as “credible” were granted much larger awards than plaintiffs perceived as “not credible”. The presence of liability issues also played a significant role in the final outcomes. Plaintiffs were considered more credible, and received much larger awards when liability issues were absent, than when they were present.

In terms of medical expert credibility, judges appear to perceive experts as more credible overall than plaintiffs, regardless of the expert’s specific involvement in the case (i.e., plaintiff v. defense testimony), or their familiarity level with the plaintiff (i.e., treating v. forensic). However, when expert categories were combined across treating and forensic categories, treating clinicians were perceived to be significantly more credible than forensic experts. The finding that medical experts were perceived as more credible than plaintiffs may be specific to FM claims or other claims involving pain conditions absent of any definitive pathology, but may not be generalizable to medical conditions with evidence of physical pathology.

Judges awarded disability claims for FM in the vast majority of cases (92%), and awards were not granted in only 7% of cases. Therefore, regardless of the ongoing debates in the medical community on the existence or diagnostic utility of FM, the Canadian courts have tended to award FM cases with disability compensation. Plaintiff credibility was clearly paramount in this process, given the significant relationship that was found between the judge’s perception of credibility and amount of awards granted. If one outcome of perceived plaintiff credibility is amount of financial award, then credibility likely plays a role in the level of access to health care. However, it remains unclear as to what specific factors were involved in judges’ determination of credibility.

The purpose of Study 2 was to examine the deliberate portrayal of excessive disability in patients with FM, and determine if certain assessment methods can accurately identify credible versus exaggerated disability in patients with FM. Even though there were statistically significant differences in patients’ self-reported problems across experimental conditions for psychological and physical functioning, those differences were not distinct
enough to differentiate the honest from the exaggeration condition with a satisfactory level of accuracy. Some psychological and PAI test combinations of two or more scores clearly improved the hit rates for exaggerators, while keeping false positives below 10%. However, most exaggerators were not identified. These findings suggest that the majority of patients with FM can exaggerate their disability on self-report measures in a plausible manner, but that some test score combinations may be useful at differentiating some patients with FM exaggerating their disability from patients with a sincere presentation of their FM.

The most surprising finding in Study 2 was that the majority of patients did not even attempt to exaggerate disability on the effort tests. This was unusual, especially when one considers that study participants typically comply with experimental instructions to exaggerate, and that a substantial minority of chronic pain patients involved in disability claims fail effort tests (e.g., Gervais et al., 2001; Schmand et al., 1998). Perhaps the lack of poor effort was related to patients with FM simply not being comfortable exaggerating, or that they were not motivated by the potential financial gain involved in disability claims. Consideration must also be given to the fact that the FM patients were not specifically instructed to exaggerate cognitive impairment.

Two very important findings emerged from Study 2. First, false positives on the Personality Assessment Inventory (PAI) validity scales and the effort tests were very uncommon. This provides the clinician some comfort that he or she will not falsely label a patient with FM as exaggerating symptoms or giving poor effort on these scales and tests. Second, the validity scales on the PAI and the effort tests failed to detect the majority of patients with FM instructed to exaggerate disability. This obviously tempers our confidence in these measures and indicates clearly that patients with FM can exaggerate disability without being detected.

The results of this study suggest that in many cases it would be extremely difficult, if not impossible, for a physician or mental health professional to determine whether a patient with FM is exaggerating their condition, or whether their presentation is sincere. If determining the sincerity of a patient’s presentation is a complex matter for the physician or healthcare professional familiar with the individual, then how accurate can a judge’s perceptions be on matters of credibility? This is an especially important question when the scientific methods aimed at these issues are still fraught with uncertainty and inaccuracies. It
is not surprising that the majority of plaintiffs (63%) in Study 1 were perceived to be partially credible, indicating a level of uncertainty in a determination of credibility.

Due to its complex nature, the legal, medical, and insurance industries will continue to struggle with questions of credibility regarding FM, and patients with FM will continue to struggle with issues of legitimacy and acceptance of their syndrome. It may be that a disability evaluation can only be made in terms of the biopsychosocial model of disease and stated in terms of reasonable probabilities (Bennett, 1996). Not only is credibility crucial to the plaintiff in litigation and disability claims situations, but as Craig and Badali (2004) point out, health care practitioners require credible and reliable information in order to assist their patients. Further research is needed from the various perspectives to assist the claimant and the courts in issues of plaintiff credibility, and to fill in the missing pieces of this profoundly puzzling and controversial disorder.
References


Credibility and Fibromyalgia


Clinical characteristics of fibrositis. I. A “blinded”, controlled study of symptoms and
 tender points. *Arthritis and Rheumatism*, 26, 817-824.

Cameron, R. S. (1995). The cost of long-term disability due to fibromyalgia, chronic fatigue
 syndrome, and repetitive strain injury: The private insurance perspective. *Journal of
Musculoskeletal Pain*, 3, 169-172.

minority differences in the pain experience of people with spinal cord injury. *Archives
of Physical Medicine Rehabilitation*, 85, 1774-1781.

education society of British Columbia.


United States and New Zealand: psychosocial and economic factors affecting severity

Assessment Inventory (PAI) and the detection of defensiveness. *Assessment*, 2, 333-
342.

Cassano, P. & Fava, M. (2004). Tolerability issues during long-term treatment with anti-

patients with fibromyalgia. *Arthritis Care and Research*, 1, 85-98.

Cavanaugh, M. E. P. (1999). Misrepresentation and non-disclosure on applications for

role of 5-HT1A and 5-HT2A receptors in depression. *Review of Psychiatric
Neuroscience*, 29, 252-265.


Crofford, L. J. & Clauw, D. J. (2002). Fibromyalgia: where are we a decade after the american college of rheumatology classification criteria were developed? *Arthritis and Rheumatism, 46*, 1136-1138.


Hadler, N. M. (1996). If you have to prove you are ill, you can’t get well: the object lesson of fibromyalgia. *Spine, 21*, 2397-2400.


Iverson, G. L. (unpublished manuscript). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering.


Iverson, G. I., Le Page, J., Koehler, B. E., Shoijania, K., & Badii, M. (under review). TOMM scores are not affected by pain or depression in patients with fibromyalgia.


impairment in healthy elderly subjects: Relationship to cortisol reactivity. *Journal of Clinical Endocrinology & Metabolism*, 82, 2070-2075.


Plesh, O., Wolfe, F., & Lane, N. (1996). The relationship between fibromyalgia and
temporomandibular disorders: prevalence and symptom severity. *Journal of 

disorders. In: F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: the fourth

illness and chronic low-back pain. The mind and the spine- which goes first? *Spine, 18*,
66-71.

Posener, J. A., DeBattista, C., Williams, G. H., Kraemer, H. C., Kalehzan, B. M., &
Schatzberg, A. F. (2000). 24-h monitoring of cortisol and corticotropin secretion in
psychotic and nonpsychotic major depression. *Archives of General Psychiatry, 57*, 755-
760.

Powell, M. R., Gfeller, J. D., Hendricks, B. L., & Sharland, M. (2004). Detecting symptom-
and test-coached simulators with the Test of Memory Malingering. *Archives of Clinical 
Neuropsychology, 19*, 693-702.

psychopathic personality features and malingering symptoms of major mental illness.

Prescott, E., Kjoller, M., Jacobsen, S., Bulow, P. M., Dansnesiold-Samsoe, B., & Kamper-
Jorgensen, F. (1993). Fibromyalgia in the adult Danish population: I. a prevalence 


antagonist dextromethorphan selectivity reduces temporal summation of second pain in


meeting of Symptom, Diagnostic and Disability Validity: Improving patient Outcomes, Markham, Ontario.


Van Houdenhove, B., Neerinckx, E., Lysen, R., Vertommen, H., Van Houdenhove, L.,


### Appendix A

#### Study 1: Variable Coding Sheet

<table>
<thead>
<tr>
<th>Case #:</th>
<th>Plaintiff/Defense:</th>
</tr>
</thead>
</table>

#### Demographic Variables
1. Gender of plaintiff:
2. Age of plaintiff (at onset of trial):
3. Time of trial (completion year):
4. Location of trial (province):
5. Residence of claimant (province):
6. Residence of claimant (city/town):
7. Marital status:
8. Duration of current marital status:
9. Number of children:
10. Years of education:
11. Employment status:

#### General Case Variables
1. Plaintiff identity:
2. Defense identity:
3. Total financial award sought:
4. Total financial award granted:
5. Award status (undetermined): 1 = no award evaluated 2 = award evaluated

#### Specific Plaintiff Variables
1. Professional who initially diagnosed FM:
2. Professional providing confirmatory/secondary diagnosis (if applicable):
3. Was confirmatory/secondary diagnosis in agreement with first? Yes No
4. Age when first diagnosis:
5. Age at onset of symptoms:
6. Is FM primary/not preceded by injury/illness, or secondary/preceded by injury/illness:
7. If secondary FM, what was primary injury/illness?
8. List co-existing medical diagnosis(es) included in claim:
9. List co-existing medical diagnosis(es) not included in claim:
10. Any psychiatric/psychological diagnosis(es) included in claim: Yes No
   Specify:
11. Any treatment for psychiatric/psychological problems included in claim: Yes No
12. Any current psychiatric/psychological diagnosis(es) not included in claim: Yes No
   Specify:
13. Any treatment for psychiatric/psychological problems not included in claim: Yes No

#### Historical Variables
1. Treatment history for current claim (circle all reported tx):
   - conventional medicine (GP)
   - conventional medicine (specialist)
   - naturopath/herbalist/chinese massage
   - chiropractor
   - PT
   - OT
   - acupuncture
   - psychiatrist/psychologist
   - pain clinic
   - support group
   - self-directed
   - other:
2. Any aches/pain complaints prior to incident/injury precipitating current claim: Yes No
3. Were the complaints from #2 related to a prior claim: Yes No
4) Any history of psychiatric/psychological diagnosis(es) not included in claim: Yes No

Specify:

Insurance/Legal Variables
1) Surveillance techniques used:
2) Defendant/insurer damages granted:
   i) contributory negligence Yes No
   ii) failure to mitigate Yes No
3) Plaintiff/insured damages (enter amounts below): Requested Granted
   i) loss of past income:
   ii) loss of future income:
   iii) earning capacity:
   iv) special damages:
   v) punitive damages :
   vi) total non-pecuniary/general:
   vii) cost of future care:
   viii) other (specify) :
4) Misrepresentation/non-disclosure/fraud issues granted: Yes No

Judge's Perceptions of Credibility
[2 = complete agreement; 1 = partial agreement; 0 = complete disagreement]
1) Plaintiff credibility:
2) Surveillance/investigative evidence credibility:

<table>
<thead>
<tr>
<th>Type of Expert</th>
<th>Expert Evidence / Plaintiff</th>
<th>Expert Evidence / Defense</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiarity to Plaintiff n</td>
<td># Experts with Credibility Rating</td>
</tr>
<tr>
<td>GP/family doctor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rheumatologist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical medicine/physiatrist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthopedic specialist/surgeon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sports medicine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurologist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurosurgeon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatrist/psychologist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counsellor/social worker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naturopath</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiotherapist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupational Therapist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiropractor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Familiarity to plaintiff, F = forensic expert/assessment for court purposes only; C = treating clinician; C & F = mixed. Credibility Rating (as perceived by the judge): 2 = complete agreement; 1 = partial agreement; 0 = complete disagreement.
Appendix B
Study 2 – Post-Experiment Interview
(Exaggerate Disability Condition Only)

For this session, you were asked to exaggerate disability.

1. Did you try to exaggerate disability? Yes No

2. Did you have problems trying to exaggerate disability?

3. In general, what was your strategy for exaggerating disability?

4. You completed questionnaires about pain, physical functioning, sleep, depression, and other problems. What was your strategy for exaggerating disability on these questionnaires?

5. On a scale from 1 – 10, how well do you think you did at exaggerating your problems? 1 means that you are not confident at all that you exaggerated in a convincing way. 10 means you are totally confident that you exaggerated in a convincing way. (Show them the scale and record their rating and comments)

6. You took some tests with the research assistant. One was a word memory test, one involved memory for pictures, and the other was on the computer. What was your strategy for exaggerating disability on these tests?

7. On a scale from 1 – 10, how well do you think you did at exaggerating your problems? 1 means that you are not confident at all that you exaggerated in a convincing way. 10 means you are totally confident that you exaggerated in a convincing way. (Show them the scale and record their rating and comments)
Appendix C
HEALTH INSURANCE INVENTORY

Subject #: ___________ Date: ___________

The following questions relate to your current knowledge of health insurance policies. Please answer all questions to the best of your ability. If you are not sure, it is okay to guess. Please indicate if you are guessing.

1. Have you ever had a life, disability, or health insurance policy through work or taken out privately? (Yes/No) If so, please state which type.

2. Have you ever filed a claim for a life, disability, or health insurance policy? (Yes/No) If so, has the claim been resolved, or is it currently in progress?

3. If you took out an individual life, disability, or health insurance policy, what past health-related information are you obligated to tell your insurance agent (e.g., whether you are a smoker, have heart problems, hypertension, diabetes, etc)?

4. Assuming you have a life, disability, or health insurance policy, what changes in your health if any, are you obligated to tell your insurance agent?

5. What do you think would happen to your life, disability, or health insurance coverage if you forgot to inform your insurance agent about any changes in your health (e.g., a doctor diagnosed you with heart problems, diabetes, rheumatoid arthritis, etc)?

6. What do you think would happen to your life, disability, or health insurance policy if you exaggerated your medical symptoms to your insurance agent?

7. Have you ever been involved in any sort of legal claim relating to a medical or health problem (Yes/No)? If so, is the legal claim ongoing? Was your Fibromyalgia included in the claim?

8. Have you ever filed a worker's compensation claim for a medical problem (Yes/No)? If so, is the claim ongoing? Was your Fibromyalgia included in the claim?