

**Complex Illness – variation and causality in
persistent medically unexplained symptoms**

Christopher D Burton

Submitted for the degree of MD

University of Edinburgh

2006



DECLARATION

I declare that I composed this thesis and initiated and conducted the research described. I was supervised in this research and received assistance with data collection. This work has been submitted only for the degree of MD.

Total word count 54,166

Christopher David Burton

Date

COPYRIGHT

Copyright in the text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the author and lodged in the Main Library at the University of Edinburgh. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Edinburgh, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

ACKNOWLEDGMENT

I wish to thank my supervisors, Professors David Weller and Michael Sharpe for their support and wisdom. I am grateful to my partners and colleagues at Sanquhar Health Centre who encouraged me to pursue the challenge of academic work while tolerating my absence and distraction.

For the collection and securing of the data for this work, I am indebted both to the participants who completed the diaries so assiduously and to Ruth Hanson, the practice research nurse, who supported them through their diary data collection. Many other colleagues both in Edinburgh and elsewhere have helped along the way to this thesis and I wish to thank them, particularly Frances Griffiths of the University of Warwick and Rachel Heath, currently of the Australian Institute of Criminology, for their advice and encouragement on the learning curve of complexity science.

Most of all, I am grateful to Diana, Lindsey, Philip and Lois, for their enduring patience and support.

This work was funded by the Chief Scientist Office of NHS Scotland through a Research Training Fellowship.

THE AUTHOR

Chris Burton graduated MBChB from the University of Edinburgh in 1982 and after GP training in Edinburgh moved to practise in Sanquhar, Dumfriesshire where he has been a GP for 20 years. He has been involved in primary care research, on and off, for fifteen years, through research networks, the CSO Research Practice Scheme and, most recently, through a CSO research training fellowship with the General Practice Section

of the School of Clinical Sciences and Community Health of the University of Edinburgh. He became a fellow of the Royal College of General Practitioners in 2006.

Table of Contents

SUMMARY	13
CHAPTER 1 INTRODUCTION	16
RESEARCH QUESTIONS	24
CHAPTER 2 MEDICALLY UNEXPLAINED SYMPTOMS AND PRIMARY CARE	25
INTRODUCTION:	25
PREVALENCE STUDIES OF MUS	27
PSYCHOLOGICAL PROCESSES IN PATIENTS WITH MUS	32
TREATMENT	37
CONCLUSIONS	40
CHAPTER 3 ELECTRONIC DIARIES FOR SYMPTOMS RESEARCH	41
INTRODUCTION	41
SYSTEMATIC REVIEW AIMS	43
METHODS	43
RESULTS	44
DISCUSSION	54
CONCLUSION.....	56
CHAPTER 4 INTRODUCTION TO THE ANALYSIS OF TIME SERIES DATA.....	57
TREND, PERIODICITY AND AUTOCORRELATION	58
MODELLING TIME SERIES, ARMA MODELS	59
MULTILEVEL MODELLING AND AUTOCORRELATION.....	60
ANALYSING TIME SERIES IN THE FREQUENCY DOMAIN.....	60
THE PROBLEMS OF MULTIVARIATE TIME SERIES.....	60
IMPLYING CAUSALITY FROM TIME SERIES DATA.....	61
COMPARING INDIVIDUAL TIME SERIES	62
CONCLUSION.....	63
CHAPTER 5 COMPLEX SYSTEMS AND NON-LINEAR DYNAMIC MODELS IN MEDICINE.....	64
STATISTICAL METHODS FOR COMPLEX SYSTEMS	67

STATISTICAL ENTROPY	68
ENTROPY, PREDICTABILITY AND HEALTH.....	69
COMPLEXITY STATISTICS IN MENTAL HEALTH	70
CONCLUSIONS	72
CHAPTER 6 AIMS AND HYPOTHESES	73
OVERALL AIM	73
SPECIFIC QUESTIONS AND HYPOTHESES.....	73
CHAPTER 7 METHODS – PATIENTS AND PROCEDURES.....	75
SAMPLE.....	75
RECRUITMENT	76
LOCATION.....	77
ETHICAL APPROVAL	77
STUDY INTERVIEWS	77
CHAPTER 8 METHODS – ELECTRONIC DIARY	81
AIMS IN DESIGN	81
HARDWARE.....	81
SOFTWARE	82
CHOICE OF PHRASES.....	83
ENTRY PROMPTS AND DATA STORAGE	84
PILOTING THE ELECTRONIC DIARY	84
REVIEW OF THE DIARY DURING THE STUDY.....	84
CHAPTER 9 METHODS – PRINCIPAL METHODS OF DATA HANDLING AND ANALYSIS	86
INTRODUCTION	86
DATA PREPARATION	86
STATISTICAL SOFTWARE USED FOR ANALYSIS.....	87
INTRA SUBJECT MEASURES.....	87
POOLING BETWEEN SUBJECTS	89
NON-LINEAR ANALYSIS	90
CHAPTER 10 RESULTS - DESCRIPTION OF PARTICIPANTS.....	93
RECRUITMENT PROCESS.....	93
EXEMPTIONS	93
PARTICIPANTS.....	94
BASELINE MEASURES	97
CONCLUSION.....	98

CHAPTER 11 RESULTS - QUALITY OF ELECTRONIC DIARY DATA.....	99
PERFORMANCE OF THE ELECTRONIC DIARY	99
ADHERENCE TO DATA ENTRY SCHEDULE	99
TEST -RETEST RELIABILITY	107
COMBINING MEASURES OF ADHERENCE AND RELIABILITY INTO A SINGLE MEASURE.	109
REACTIVITY	110
CONCLUSION: PERFORMANCE OF ELECTRONIC DIARIES	115
CHAPTER 12 RESULTS - PRELIMINARY TIME SERIES ANALYSIS.....	116
INTRODUCTION	116
CHOICE OF DATA.....	116
DESCRIBING THE NATURE OF THE TIME SERIES DATA	117
RESULTS OF AUTOCORRELATION FUNCTION.....	117
DIFFERENCES BETWEEN SYMPTOMS IN MEASURES OF AUTOREGRESSION.....	118
EFFECT OF DIFFERENCING THE DATA.	119
TESTING FOR NORMALITY OF DISTRIBUTIONS.	120
CONCLUSION: PRELIMINARY TIME SERIES ANALYSIS.....	122
CHAPTER 13 RESULTS - CORRELATION WITHIN SUBJECTS.....	123
INTRODUCTION	123
GENERATION OF CORRELATION MATRICES.....	123
POOLING OF CORRELATION COEFFICIENTS BY META-ANALYSIS OF REGRESSION COEFFICIENTS	129
HETEROGENEITY OF CORRELATION COEFFICIENTS.....	135
MAGNITUDE OF THE CORRELATION COEFFICIENTS.....	136
CONCLUSION.....	138
CHAPTER 14 RESULTS - TEMPORAL SEQUENCES.....	139
INTRODUCTION	139
GRANGER CAUSALITY	139
RESULTS OF GRANGER CAUSALITY TEST	140
CONCLUSION.....	145
CHAPTER 15 RESULTS - CHANGES RELATED TO CONSULTATIONS.....	147
INTRODUCTION	147
DATA AND METHODS	147
COMPARISON OF BASELINE AND PRE-CONSULTATION	147
COMPARISON OF PRE- AND POST-CONSULTATION	149
CONCLUSION.....	151
CHAPTER 16 RESULTS - ANALYSIS OF COMPLEXITY.....	152

INTRODUCTION	152
SAMPLE ENTROPY METHOD.....	152
GENERATION OF SURROGATE TIME SERIES AND SAMPLE ENTROPY DISTRIBUTIONS.....	152
DISTRIBUTIONS OF SAMPLE ENTROPY RESULTS.....	153
RANKING OF ACTUAL SAMPLE ENTROPY WITHIN SURROGATE DISTRIBUTIONS	156
CONVERSION OF SAMPLE ENTROPY & RANKS TO Z SCORES	157
Z SCORES	159
SENSITIVITY OF SAMPLE ENTROPY TO LARGE SCALE PROPERTIES OF THE DATA.....	161
SENSITIVITY OF SAMPLE ENTROPY TO DATA QUALITY.....	162
SAMPLE ENTROPY DIFFERENCES BETWEEN VARIABLES.....	162
COMPARISON WITH BASELINE VARIABLES	163
CLUSTERING BY SAMPLE ENTROPY RESULTS.....	163
CONCLUSIONS.....	165
CHAPTER 17 RESULTS - INTERVIEWS AND QUALITATIVE DATA.....	166
INTRODUCTION	166
USEFULNESS OF THE DIARY.....	166
CHOICE OF THEMES.....	167
THE STRONG INDIVIDUAL.....	168
THE CHAOTIC NARRATIVE.....	169
AMBIVALENCE OVER LABELS.....	171
ALL OR NOTHING BEHAVIOUR.....	174
LINKING MIND AND BODY IN EXPLANATORY MODELS.....	175
CONCLUSION.....	177
CHAPTER 18 DISCUSSION – REVIEW OF MAIN FINDINGS	178
SPECIFIC QUESTIONS AND HYPOTHESES.....	178
CHAPTER 19 DISCUSSION – STRENGTHS AND LIMITATIONS OF THE STUDY	181
DIARIES	183
ANALYSIS	185
QUALITATIVE ANALYSIS	187
CHAPTER 20 DISCUSSION – ANXIETY, DEPRESSION, CONCERN AND FUNCTIONAL PHYSICAL SYMPTOMS.....	188
CURRENT MOOD – THE EFFECT OF LOW MOOD AND DEPRESSION.....	188
STRESS, ANXIETY AND CONCERN	194
CHAPTER 21 DISCUSSION – VARIATION, CAUSALITY AND MEANING	201
OBSERVATION AND CAUSALITY.....	201

SOCIAL CONSTRUCTION OF MEANING IN MUS	203
A CRITICAL REALIST APPROACH.....	204
IMPLICATIONS OF A CRITICAL REALIST APPROACH TO MUS.....	207
CHAPTER 22 DISCUSSION - TOWARDS A MODEL OF MUS AS COMPLEX ADAPTIVE ILLNESS.....	211
CHAPTER 23 CONCLUSIONS.....	221

List of Appendices

APPENDIX A STUDY PROTOCOL.....	A-1
APPENDIX B CONSENT FORM	B-1
APPENDIX C HAD SCALE	C-1
APPENDIX D WHITELY – 7 ILLNESS CONCERN QUESTIONNAIRE.....	D-1
APPENDIX E SOMATIC SYMPTOM INVENTORY	E-1
APPENDIX F SELF REFERRAL GP CHECKLIST	F-1
APPENDIX G INSTRUCTIONS FOR HANDHELD COMPUTER DIARY	G-2
APPENDIX H META-ANALYSIS COMPUTER SCRIPT (PYTHON).....	H-1
APPENDIX I GRANGER CAUSALITY COMPUTER SCRIPT (R)	I-1
APPENDIX J SAMPLE ENTROPY COMPUTER SCRIPT (PYTHON)	J-1
APPENDIX K INCLUDED PAPERS	1

List of Tables

TABLE 1 PREVALENCE OF MEDICALLY UNEXPLAINED SYMPTOMS AS THE MAIN REASON FOR CONSULTING A GP.....	29
TABLE 2 OVERLAP OF SOMATISATION AND PSYCHIATRIC DISORDERS IN PRIMARY CARE / COMMUNITY SAMPLES	30
TABLE 3 ELECTRONIC DIARY STUDIES OF ASSOCIATIONS BETWEEN PHYSICAL SYMPTOMS AND PSYCHOSOCIAL VARIABLES	46
TABLE 4 STUDIES EVALUATING ELECTRONIC DIARIES SOLELY IN TERMS OF ACCEPTABILITY ACCURACY OR COMPARED TO PEN AND PAPER DIARIES	49
TABLE 5 KEY FINDINGS OF DIARY STUDIES.....	52
TABLE 6 CHARACTERISTICS OF NON-LINEAR DYNAMIC SYSTEMS	65
TABLE 7 QUESTIONS ASKED IN SYMPTOM DIARY	84
TABLE 8 BREAKDOWN OF PARTICIPANTS BY AGE, GENDER, PRINCIPAL SYMPTOM AND OCCUPATIONAL STATUS.....	95
TABLE 9 SYMPTOMS RECORDED BY EACH PARTICIPANT	96
TABLE 10 DATA ENTRY RATES FOR EACH PARTICIPANT.....	104
TABLE 11 INTRAClass CORRELATION COEFFICIENT FOR VARIABLES	107
TABLE 12 INTRAClass CORRELATION COEFFICIENT BETWEEN INDIVIDUALS	108
TABLE 13 DIURNAL VARIATION IN SYMPTOMS:	114
TABLE 14 META-ANALYTIC CORRELATION COEFFICIENTS FOR VARIABLE PAIRS	130
TABLE 15 COMPARISON OF META-ANALYTIC CORRELATION COEFFICIENTS BETWEEN HIGH AND LOW SCORES ON BASELINE SCALES.....	136
TABLE 16 META-ANALYTIC CORRELATION COEFFICIENTS SHOWING THE LARGEST DIFFERENCES BETWEEN HIGH AND LOW SCORERS ON BASELINE SCALES	137
TABLE 17 DISTRIBUTION OF INTERACTIONS YIELDING HIGH GRANGER SCORES CROSS TABULATION OF THE HIGHEST 10% OF SCORES.....	145
TABLE 18 PAIRED VALUE T TESTS COMPARING PRE-CONSULTATION WITH BASELINE.....	148
TABLE 19 PAIRED VALUE T TESTS COMPARING PRE-CONSULTATION WITH POST-CONSULTATION	149
TABLE 20 MEAN SAMPLE ENTROPY FOR DATA AND SURROGATES BY COMMON SYMPTOMS.....	156
TABLE 21 Z SCORE FOR SAMPLE ENTROPY, FOR EACH OF THE FIVE SYMPTOMS RECORDED BY EACH PARTICIPANT	159
TABLE 22 AVERAGE SAMPLE ENTROPY Z SCORE FOR PHYSICAL SYMPTOMS	160
TABLE 23 CORRELATION COEFFICIENTS BETWEEN SAMPLE ENTROPY OR Z SCORE AND LARGE SCALE CHARACTERISTICS OF TIME SERIES	161
TABLE 24 COMPARISON OF Z SCORES FOR COMMON VARIABLES IN THE	164
TABLE 25 CLASSIFICATION OF EXPLANATORY STRUCTURES.....	202

List of figures

FIGURE 1 - OVERLAP OF FUNCTIONAL SYNDROMES – AFTER DEARY	26
FIGURE 2 FLOW OF PARTICIPANTS THROUGH STUDY	94
FIGURE 3 HISTOGRAMS OF BASELINE QUESTIONNAIRE SCORES	97
FIGURE 4 EXAMPLES OF HISTOGRAMS OF ENTRY TIME	101
FIGURE 5 COMPARISON OF DATA ENTRY RATES BY QUALITATIVE SCHEDULE ADHERENCE	105
FIGURE 6 HISTOGRAM OF THE ABSOLUTE DIFFERENCE BETWEEN CHECK AND MATCHING ORIGINAL ENTRIES	106
FIGURE 7 COMPARISON OF RELIABILITY OF ENTRY WITH ADHERENCE TO DIARY SCHEDULE	109
FIGURE 8 TESTS FOR REACTIVITY: CHANGES IN MEAN AND STANDARD DEVIATION OVER TIME	111
FIGURE 9 TESTS FOR REACTIVITY IN MEAN VALUE BY INDIVIDUAL AND VARIABLE	113
FIGURE 10 AUTOCORRELATION CATEGORY OF COMMON SYMPTOMS	118
FIGURE 11 BOX PLOT OF AUTOREGRESSION COEFFICIENTS	119
FIGURE 12 TESTS FOR NORMALITY IN DIFFERENCED TIME SERIES	121
FIGURE 13. NETWORK DIAGRAM OF EACH INDIVIDUAL SUBJECT CORRELATION MATRIX	125
FIGURE 14 - EXAMPLES OF FOREST PLOTS OF META-ANALYSES OF CORRELATION COEFFICIENTS.	132
FIGURE 15 HISTOGRAM OF Z SCORES FROM GRANGER TESTS ON ALL VARIABLE PAIRS	140
FIGURE 16 DIRECTIONAL GRANGER CAUSALITY - PSYCHOLOGICAL VARIABLES	141
FIGURE 17 DIRECTIONAL GRANGER CAUSALITY – PHYSICAL SYMPTOMS	142
FIGURE 18 BREAKDOWN OF PSYCHO-SOMATIC GRANGER CAUSALITY	143
FIGURE 19 BREAKDOWN OF SOMATO-PSYCHIC GRANGER CAUSALITY	144
FIGURE 20 SCATTER PLOT OF CONCERN IN BASELINE AND PRECONSULTATION PERIOD	148
FIGURE 21 SCATTER PLOT OF STRESS IN PRE-CONSULTATION AND POST-CONSULTATION PERIOD	150
FIGURE 22 SCATTER PLOT OF ILLNESS CONCERN IN PRE-CONSULTATION AND POST-CONSULTATION PERIOD	150
FIGURE 23 HISTOGRAM OF SAMPLE ENTROPY SCORES	154
FIGURE 24 SCATTER PLOT OF ACTUAL SAMPLE ENTROPY OF DATA AGAINST MEAN SAMPLE ENTROPY FROM SURROGATE DATA	155
FIGURE 25 HISTOGRAM OF RANKING OF ACTUAL SAMPLE ENTROPY WITHIN SURROGATE DISTRIBUTIONS.	157
FIGURE 26 Z SCORES FOR SAMPLE ENTROPY	158
FIGURE 27 HISTOGRAM OF Z SCORES OBTAINED BY CALCULATION	159
FIGURE 28 SCATTER PLOT OF SAMPLE ENTROPY AGAINST DATA QUALITY	162
FIGURE 29 SAMPLE ENTROPY BY ACTIVITY STATUS	163
FIGURE 30 ACTIVITY STATUS BY SAMPLE ENTROPY CLUSTER	164
FIGURE 31 REVISED MODEL OF MUS, AFTER DEARY AND FIGURE 1.	216

List of Abbreviations

ANOVA.....	Analysis of variance
ARMA.....	Autoregressive Moving Average
CBT	Cognitive Behavioural Therapy
CFS	Chronic Fatigue Syndrome
EEG	Electroencephalogram
FSS	Functional Somatic Syndrome
FM	Fibromyalgia
GHQ	General Health Questionnaire
GP	General Practitioner
HAD	Hospital Anxiety and Depression Scale (-A, anxiety subscale; -D depression subscale)
IBS	Irritable Bowel Syndrome
ICC	Intra-class Correlation Coefficient
ICQ	Illness Concern Questionnaire
MCS	Multiple Chemical Sensitivity
MLM	Multi-level Modelling
PDA	Personal Digital Assistant
RCT	Randomised Controlled Trial
SSI	Somatic Symptoms Inventory
SSRI	Selective Serotonin Reuptake Inhibitor
TMD	Temporomandibular (joint) Dysfunction
VAS	Visual Analogue Scale

Summary

Introduction

Some patients' experience of illness cannot be adequately explained by medicine's current knowledge of disease and the term Medically Unexplained Symptoms (MUS) is commonly used to describe illness where symptoms cannot be explained by organic pathology. Epidemiology shows that experience of MUS is associated with psychological distress and a prevailing attitude among professionals has been that MUS represents the somatisation, or bodily experience as physical illness, of mental distress. This thesis challenges the idea that somatisation is the predominant mechanism by investigating the links between physical symptoms and psychological states dynamically using quantitative symptom diaries.

Research Questions

The thesis uses diary and interview data to address five research questions:

How do symptoms and emotional states reported by electronic diary vary over time in patients with persistent MUS?

What are the concurrent associations between symptoms and mood?

Is there any evidence for consistent sequential relationships between symptoms over time?

Do symptom time series data for patients with MUS show signs of loss of complexity?

How do patients with MUS describe their condition after viewing their own diary data?

Methods

Twenty six participants, each of whom had at least three apparently functional physical symptoms, completed diaries twice daily for up to 12 weeks using electronic diaries on handheld personal computers. Each diary entry comprised eight visual analogue scales representing three personally relevant symptoms, fatigue, mood, anxiety, illness concern and stress. Brief psychometric scales for depression, anxiety, illness worry and somatisation were carried out at enrolment and a longer

unstructured interview was carried out after the diary data collection at which the participants' results were fed back to them.

Results

The diaries were acceptable to participants and generated good compliance and data with high reliability. The data showed considerable day to day variation for most participants with strong short term autocorrelation within the time series data.

Correlations were calculated between each pair of variables for each individual, after adjusting for autocorrelation, and individual participant data were pooled by meta-analysis. There was heterogeneity in the correlations between variable pairs, and between individuals recording the same variable pair. Correlations were strongest for painful symptom severity and low mood ($r = 0.15-0.4$); correlations between physical symptoms and anxiety or stress were generally weak or insignificant.

Sequential relationships between variables were tested using granger causality which estimates the influence of lagged (prior) values of one variable on the prediction of another by regression. There was significant heterogeneity in the strength of granger causality, in particular visceral pain and headache had greater sequential associations with low mood than musculoskeletal pain despite the fact that musculoskeletal pain had stronger concurrent correlation with mood.

While concern about symptoms was associated with their severity, particularly for abdominal pain, it did not show significant changes before or after consultations with the participant's GP which occurred during the diary collection period.

Sample entropy, a measure of complexity within the time series data, was significantly reduced compared to controls, in support of the hypothesis that MUS would be associated with reduced statistical complexity.

Interview data confirmed the presence of chaotic narratives and a struggle for meaning and coherence through the symptom experience which contradicted conventional simple medical explanations.

Conclusions

Patients' recorded experiences refuted simple associations between day to day stress and current symptoms, however changes in mood were more strongly associated with changes in symptoms. The study is the first to apply a measure of complexity to

physical symptom diaries and shows it to be reduced in keeping with a “loss of complexity” hypothesis. The collection of diary data appeared to encourage constructive reflection on the contents and participants saw the diaries as worthwhile. They were generally open to explanations linking mind and body, particularly when these drew on positive attributes. Simple stress induced illness models were neither popular nor supported by the data. Instead there is a need for explanatory models which acknowledge the complexity of the experience of persistent medically unexplained symptoms

Chapter 1 Introduction

Some patients' experience of illness cannot be adequately explained by medicine's current knowledge of disease. Distressing physical symptoms and their associated disability often appear to arise from physiological and neuro-psychological processes which are largely within, or contiguous with, the normal range for healthy individuals. It has traditionally been suggested that these disorders represent a physical representation of emotional distress, but more recent developments in understanding have highlighted complex but subtle interactions between mind and body, and between patient and doctor. This thesis addresses quantitative measures of variation and notions of causality, in a sample of patients with so-called medically unexplained symptoms, to explore the ways in which these are complex illnesses.

This introductory section sets the background to the thesis by briefly summarising current theories about how patients experience symptoms and decide to consult a doctor. It then introduces two themes which are reviewed in detail in subsequent chapters (medically unexplained symptoms, electronic diaries for symptoms research) and two which are introduced in sufficient detail to permit understanding of the methods and analysis of the research project (complexity science in the understanding of illness models and an introduction to time series analysis), before concluding by outlining the key research questions of the thesis.

Patients are people with symptoms who consult a doctor

The primary function of general practitioners, defined in 1972(1) and retained in their current contract(2), is "the management of patients who are, or who believe themselves to be, ill". This apparently simple definition, immediately unites both physical and psychological aspects of illness. It goes beyond the limits of known organic disease and demonstrable pathology behind which much of modern medicine has retreated, to include patients' beliefs and concerns. While the phrase "believe themselves to be ill" implies an excessive degree of certainty - "believe they may be ill" is perhaps nearer the mark for day to day primary care - the principle of subjective perception of illness by the individual, as opposed to objective identification of disease by the professional, remains at the heart of clinical general practice.

Symptoms are bodily feelings that may represent illness

Cognitive models of human behaviour tell us that individuals continuously sense bodily changes and react to them. For most of us, most of the time, these changes are simply the innate variations of homeostasis or adaptation to an ever changing environment and the reactions are automatic. Sometimes however they are a possible sign of bodily disease or malfunction and require greater attention, comparison with past experience, and interpretation in the light of personal and shared knowledge.

Models such as Leventhal's self regulation theory(3) outline a system of continual appraisal of sensory inputs which is broadly supported by current neuroscience(4). This appraisal appears to comprise elements of appraisal, comparison and interpretation in which sensations are matched with previous experience and knowledge of what they might mean. Frequently, this results in normalisation – the recognition that changes in the way one feels are not representative of anything other than homeostasis and the environment – but this is not always the case. In some circumstances, sensations are interpreted as potentially due to illness and bodily sensations become symptoms.

The process of making sense of symptoms and possible illness has been extensively evaluated(5) in the context of lay models of illness. These suggest that illness beliefs requires an identity for the condition, a cause (both general and personal), a timescale and a natural history, as well as understanding of how to deal with it.

In some cases the conversion of sensation to symptom to illness experience is straightforward. For instance, the sensations of a stuffy nose and itchy eyes after walking through a field of long grass clearly represent symptoms of hay fever when one has been bothered by the same thing every June for years. But this is not always the case even when it seems apparent that it should be: for instance Weinman(6) and colleagues demonstrated that a substantial proportion of patients with chest pain signifying a heart attack either fail to recognise the sensation as a symptom of illness, or rationalise the illness as something other than heart disease.

Symptoms do not consistently lead to consultation

People who believe they may be ill consult a doctor for a range of reasons relating to their symptoms, these include resolution (cure), relief, clarification (diagnosis),

interpretation and support in coping. The decision to consult is a function of the severity of the symptoms, their personal relevance and manageability to the individual, and the norms and pressures of family, friends and society(7;8). Studies have consistently shown that symptoms, as possible signs of disease, are far more commonly tolerated or self- managed than they are presented to a doctor or other professional(9). This so called symptom iceberg(10) offers the possibility that different people present differing proportions of symptoms or, to continue the analogy, may have a greater or lesser proportion of the iceberg remain hidden underwater.

While personal cognitive factors play a part there is also evidence for social processes(8;9) and past experience of the healthcare system(11;12) governing consultation decisions. These decisions include both consulting and choosing not to(13). At the largest scale, a national healthcare system, I have shown elsewhere(14) that the distribution of consultation patterns is best explained by a model suggesting a complex adaptive social system with societal “rules” about how and when to consult.

Studies of the effect of healthcare practice on further consultation show that giving a prescription for a self-limiting condition increases the probability of further consultations(15), and conversely that giving patients explicit information, which is both patient centred and acknowledges the limitations of medical treatment, can reduce it(16).

Not all symptoms are explained by disease

If the boundary between sensation and symptom is blurred at one (severe) end of the spectrum, one would expect this to be the case at the other, milder, end of the range. Several studies show that most people experience physical sensations that might be symptoms of illness most days, but generally deal with them, usually by normalisation. Even when sensations are seen as symptoms, people usually manage them themselves, waiting for resolution or using simple interventions or medications. That way, over two thirds of patients who consult a doctor with physical symptoms have evidence of disease, whether acute or severe, chronic or degenerative, or mild and self limiting. However, in around a fifth of GP consultations(17) and between a third(18) and a half(19) of specialist consultations there is no demonstrable disease.

While the symptoms leading to these consultations are diverse in nature and in location within bodily organ systems, they appear to share common patterns of illness experience and behaviour and are known as medically unexplained symptoms (MUS) or functional somatic symptoms (FSS) (20).

Historically MUS have been seen as a manifestation of somatisation, the process by which emotional distress is experienced as, or presented to others including physicians in the form of physical symptoms(21). This primacy of the psychological is also included in the commonly used term “psychosomatic” which is so often resisted by patients(22). While there is undoubtedly a link between symptoms and psychological distress this is not limited to symptoms unexplained by organic disease and not all patients with MUS have psychological illness. The literature pertaining to MUS in primary care will be reviewed in detail in chapter 2.

The problem of finding a cause for the “unexplained”

While “medically explained” symptoms have, obviously, medical explanations which are broadly shared by (most) physicians and patients, the cause of medically unexplained symptoms is inevitably in greater doubt.

In a simplistic approach, attempts to explain the medically unexplained can tackle the problem at any of three levels: the physiological processes which give rise to physical sensations, the cognitive processes of converting sensation to symptoms, and the interpretation of symptoms as illness.

In a few patients, the processes in the second and third stages of this sequence are so abnormal that they merit psychiatric diagnoses such as somatisation disorder(23) and hypochondriasis(24). A further uncommon group comprises patients with severe functional impairment meriting the label of conversion disorder. In most cases however the cognitive and interpretive patterns are not so obviously different from patients with explained symptoms that clear distinctions can be drawn.

In other situations common sense inferences can be made and used, for instance a person anxious about heart disease is more likely to label the feeling of acid in his oesophagus as the symptom of chest pain and interpret it as possibly indicating angina. Likewise, muscle aches and fatigue may simply be the consequences of doing too much, or not sleeping well because of pressing worries. However these are

only partial explanations, and are not good at explaining why there are inconsistent associations between symptoms and possible cause in the individual.

Epidemiologically, there is strong evidence for an effect of current mood (such as depression) on symptoms(9) and also for past events, including illness experience and abuse.

While common sense and epidemiological explanations hold a degree of validity, they also frequently carry moral implications(25), of personal weakness, or of not trying hard enough to cope. Such implications are frequently at odds with individuals' premorbid personalities and frequently lead to isolation and resentment as the validity of patients' personal experience is called into question(26;27).

Given the weakness and the lack of specificity of the cognitive explanations, and the implicit moral dimension to their common use, it is unsurprising that patients and doctors frequently search for new or missed physical disease explanations for the unexplained, sometimes leading to new "epidemics" of a condition such as myalgic encephalomyelitis.

The problem of probabilistic inference versus personal cause

The nature of scientific enquiry contrasts starkly with that of personal experience. Outside of pure sciences such as physics, most cause is assumed to be complex and probabilistic, with causal factors rather than direct causes and an additional component of "error" representing either systematic or random variations and chance(28). This recognition underpins much of our understanding of disease obtained through disciplines such as epidemiology. Nonetheless society, and medicine, still has a hankering for certainty rather than probability.

Probabilistic science deals with its uncertainty by sampling from multiple cases and then inferring statistically between them. By balancing out as much error as possible, it seeks to identify common processes. Personal experience is different. It follows a narrative trajectory which uses its own time scales(29) of point events rather than chronology and it is confined to the one person who both experiences and interprets it. Except where separate episodes of similar symptoms can be seen as sufficiently similar to make comparisons, probabilistic analysis across a sample of instances is impossible, and so every episode is in some way unique. This uniqueness of personal

experience, and its contrast with the generality of medical knowledge is one reason why communication between doctors and patients sometimes breaks down(30).

Psychologists recognise however that narrative memory is not constant and is influenced both by an individual's current state, such as their mood or arousal and by intervening events between the time of the object being recalled and the present(31;32). These recall biases mean that personal narrative memory is effectively a story with a purpose, both to describe events but also to make sense of them(33). As such it fulfils a greater purpose, both personally and socially, than simply a literal record of events.

Explanations are important

Cognitive models such as those mentioned show that patients make sense of their explanations in a fairly consistent way. These explanations commonly draw on both folk illness models(5) and on biomedical ideas and terms(34;35) and are frequently complex.

Simple reassurance of the absence of pathology has been found to be only weakly effective for patients(36-38) with medically unexplained symptoms and it appears that in other situations phrases intended to reassure actually have the opposite effect(39).

Studying patients with medically unexplained symptoms in both secondary and primary care, Salmon and colleagues identified a mismatch between the explanations most patients received and those which were actually perceived as helpful(40). They postulated that the doctors and patients' expectations of medical encounters were sufficiently different that professionals seeking to provide what they mistakenly thought patients wanted(41) (a further definitive medical opinion to resolve the doctor's uncertainty) failed to meet the needs of their patients. More constructively they also identified a model whereby normalisation may be appropriate and proposed ways in which it may be effectively carried out(42).

Instead of making coherent explanations of symptom disorders as illnesses, much work over the last 25 years has been concerned with the notion that medically unexplained symptoms represent a method of communicating psychological distress which is either not recognised or which cannot be articulated(21). This has led to

treatment models geared to the reattribution of somatic symptoms to psychological disorders and processes.

However while much effort has gone into moving physical symptoms towards psychological disturbance it is important to note that Goldberg and Bridges(43), in raising the profile of the process of somatisation, indicate that the adoption of the illness role with physical symptoms is a socially acceptable way of acknowledging incapacity and that by offering an escape from social pressures may enhance mental health by preventing depression. With hindsight some of their enthusiasm for reattribution may have been influenced by a more optimistic view of medicine's ability to cure mental distress arising from difficult lives(44).

Towards a coherent explanation

An alternative to reattribution may be to consider symptom disorders as illnesses in their own right which have both physiological and cognitive processes, both of which are amenable to analysis and to therapeutic intervention and both of which need to be considered in order to fully give meaning to the illness to the individual(35;45).

Such an explanation requires particular focus on the nature of cause, for it must balance both personal experiential knowledge(30) against medical scientific method and integrate the psychological and the physiological components of the illness. Given the documented variation in symptom disorders over both long and short term time, any explanatory model must be capable of generating complex changes. In contrast, the relatively simple analogies used in reattribution(46), (for instance the analogy that as arm muscular tension from lifting causes pain, so cranial muscle tension will cause headache), while tangible, are too easily refuted from experience,. Many such explanations which lack predictable recurrence also suffer from a lack of objective scientific evidence to back them up.

An alternative to a causal system of interactions

Conventional statistical models of cause and effect posit causal relationships between variables which are then obscured, to some greater or lesser extent, by random and systematic variation in context and measurement leading to a probabilistic model.

An alternative approach to complicated systems is to view interactions as contingent rather than consistent, being dependent on the local context rather than interfered

with by it. Theoretical models of complex adaptive systems, which use rules of local interaction in response to events and changes suggest that the capacity to adapt to an ever-changing environment derives from contingent, context specific, effects. To study this, new forms of analysis and simulation are needed, not to reveal the underlying relationship between causal factors and effects, but to examine the changes over time - the dynamics - of the system and how it responds. Such investigation, at least within the quantitative research domain, draws on the physics of non-linear and complex systems(47;48). This non-linear or complex systems approach has been demonstrated with time series data in several fields of physiology and mental health; its potential relevance to MUS will be reviewed in Chapter 5.

Approaches to analysis

Conventional psychological methods are divided between those which are idiographic, focusing on the individual, and those which are nomothetic, seeking generalisations between individuals. Usually idiographic methods are qualitative, whereas the nomothetic approach is quantitative. An approach which bridges some of the gaps between these methods is to use repeated quantitative measures at an idiographic level, thereafter making generalisations between individuals with the results of the initial analysis. In order to incorporate temporal relationships, data for such methods must represent an ordered time series.

Investigations which generate multivariate time series data based on self reported data are able to address questions of associations between variables which include a time dimension. While making causal inferences from associations requires caution to avoid making spurious assumptions it has been used in medical science and in other fields such as economics.

Time series studies have been relatively infrequent in symptoms and psychological research(49), not least due to the difficulties of accurate data collection and handling. These technical problems can now be overcome by the use of electronic diary methods(31). A detailed systematic review of the use of electronic diaries for symptoms research appears in Chapter 3.

Choosing the research questions

It is clear that research is increasing our understanding of the physiological and cognitive processes in play for patients with medically unexplained symptoms in primary care, but also that the forms of explanation that patients receive are inadequate and fail to meet their needs(40;50). This thesis concerns research which seeks to characterise the interactions between functional somatic symptoms and psychological states which occur over time using analysis of data from electronic diaries.

The research questions were chosen to examine the validity of electronic diary data and use it to identify associations between psychological and physical symptoms at the individual patient level through a range of analytical techniques. Additionally some simple observations were used to explore the value of explanations based on personal data to patients themselves.

Research Questions

The overall aim of this research is to investigate the interaction of mind and body over time in patients with persistent medically unexplained symptoms.

The specific research questions will be discussed in more detail in section 6 but are briefly listed below

Are electronic diaries a suitable tool for symptoms research?

How do symptoms and emotional states reported by electronic diary vary over time in patients with persistent MUS?

What are the concurrent associations between symptoms and mood demonstrable in individuals and how much to they vary between individuals?

Is there any evidence for consistent sequential relationships between symptoms over time?

Do symptom time series data for patients with MUS show signs of loss of complexity?

How do patients with MUS describe their condition after viewing their own diary data?

Chapter 2 Medically unexplained symptoms and primary care

Introduction:

This chapter reviews the literature on medically unexplained symptoms in relation to primary care with three aims. The first is to review the epidemiology of MUS in primary care and the overlap between MUS and psychiatric diagnoses; the second is to summarise current understanding of the psychological processes which appear important in these disorders and the third is to review treatments for MUS which are applicable or available to general practitioners. The chapter begins with a short discussion of the kinds of conditions under the umbrella term of MUS and the difficulty of classification.

A fundamental element of primary care is dealing with symptoms which may, or may not, be due to physical disease. Patients attend with specific symptoms for a variety of reasons(8) including their severity, the disruption they cause, and concerns in the patient's mind about what they may represent(51). While most people experience at least some physical symptoms, a number of patients repeatedly attend with symptoms for which a conventional pathology cannot be identified. "Symptom Syndrome"(52) clusters are widely recognised and include irritable bowel syndrome, chronic pelvic pain, and fibromyalgia. Studies of patients with these conditions have found striking similarities between them(20), with a substantial proportion of patients showing evidence of psychological distress(53) which is either not expressed or unrecognised in the medical consultation.

In an attempt to explain this process psychiatrists have used the term somatisation, although the meaning of this term has changed over time(54). Initially, it was thought of as similar to hysterical conversion. It now effectively has two meanings: the expression of psychological illness through physical symptoms(43), (as in the term "somatised depression") and repeated medical help-seeking for multiple medical symptoms without organic disease(23) (as in "Somatisation Disorder"(55)). These two concepts overlap but they are not synonymous. Bridges and Goldberg further divided somatised psychiatric disorder as partial somatisation, where the patient

acknowledges the possibility that psychological distress may be causing the symptom when directly questioned, and true somatisation where the patient does not acknowledge any psychological link when challenged, despite meeting diagnostic criteria(56).

To overcome the confusion around the term somatisation, many researchers prefer the term “Medically Unexplained Symptoms”(57). While this recognition of uncertainty is helpful in a research environment, the fact that the meaning of physical experiences seems fundamental to these conditions(58) makes it inappropriate for clinical care and it has been criticised on these grounds(21). As for alternatives, “Psychosomatic Illness” is seen in the public eye as synonymous with being “all in the mind”(22), while “Functional Somatic Symptoms”(20) may be preferable but is not in routine use. More recently Kroenke has suggested a further simplification: Physical Symptom Disorder(59). In this review I have used the term Medically Unexplained Symptoms (MUS) to mean “physical symptoms for which no clear or consistent organic pathology can be demonstrated (although organ dysfunction may be an integral part of the symptom)”.

Figure 1 - Overlap of functional syndromes – after Deary

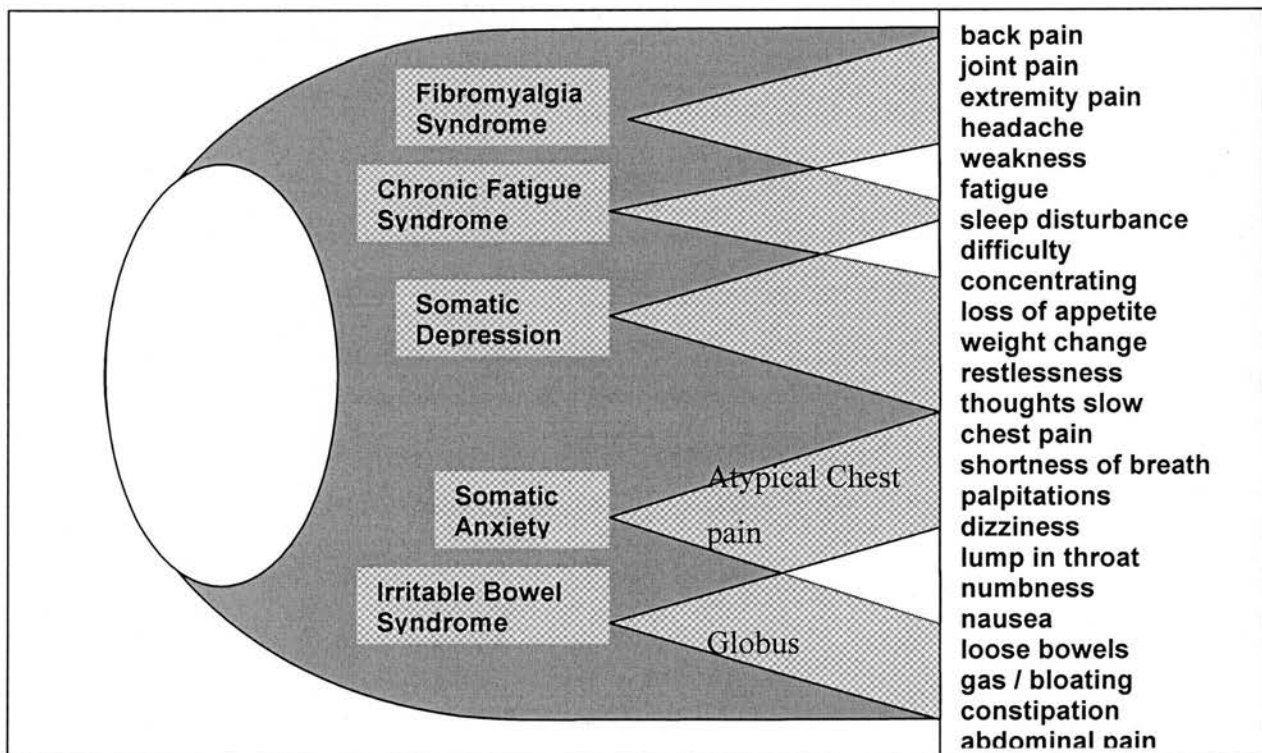


Figure 1 shows a selection of the symptoms and syndromes under review. Before considering the psychosocial elements of MUS, it is important to consider recent developments in the pathophysiology of the conditions. Increasingly, subtle physiological changes in structure(60), hypothalamo-pituitary axis function(61), neurotrophic activity(62) and inflammatory markers(63) are being recognised both in symptom disorders and in chronic pain. These remind us that current medical knowledge is far from complete and also that the boundary between “organic” and “functional” may be at least blurred, and at most artificial.

Prevalence studies of MUS

Studies of the prevalence of MUS and overlap with psychiatric illness

Studies estimating the prevalence of MUS in primary care can be grouped into two categories: those that use the main reason for the consultation to determine whether the problem is unexplained or not, and those that apply measures of somatisation to populations including community samples, primary or secondary care patients, and particular groups such as frequent attenders.

Prevalence of medically unexplained symptoms as the reason for consulting

The MEDLINE and EMBASE databases were searched from 1980 to 2006 for a combination of one or more of “medically unexplained symptoms”, “somatisation” or “somatoform disorders” with one or more of “general practice”, “family medicine”, and “primary health care”. This identified six studies of the prevalence of MUS as a reason for consulting in primary care; these are summarised in

Table 1. The UK studies of Mumford(64) and Peveler(17) identified a physical symptom without likely organic disease as the main reason for 15% and 19% of consultations respectively.

These estimates do not differentiate between consultations which are effectively dealt with and do not lead to repeat consultation, and those which lead to a more persistent problem. Recent work from the Netherlands suggests that repeated consultation for MUS is rather less common(65), with only 2.5% of patients consulting 4 or more time within a year with any unexplained symptom.

Prevalence of somatisation disorders in primary care and general populations

The search strategy identified ten studies of somatisation disorders with sample sizes of over 100 individuals from general populations or primary care consultants; these are shown in Table 2. These used a variety of criteria but all included patient self-ratings of the presence of symptoms and used cut-off points based on the number, rather than the character, of symptoms. As well as recording the prevalence of patients reporting above a set number of symptoms, most of these studies identified the prevalence of psychiatric disorder.

The results of these studies are highly dependent on the criteria used both in symptom counts and for severity of psychiatric disorder. While less than 0.5% patients meet the criteria for DSM Somatisation Disorder(55) (including at least 8 from 40 symptoms due to non-organic disease in at least four bodily systems with age of onset before 30), 16-22% meet(66-68) the abridged somatisation criteria of 4/37 symptoms for men and 6/41 for women. Over half of one sample of patients(69) admitted to at least one MUS causing some interference with their life. Similar variation in prevalence is seen with concurrent mental illness. While as few as 20% of patients with only one MUS have a current psychological illness(69), the proportion rises to over 30% with 4 symptoms(66) and over 80% with 10 or more(68), regardless of whether they are medically explained or not(70).

These studies consistently demonstrate that while MUS are common, and are often associated with psychiatric morbidity, many patients with MUS have no definite psychological illness and that patients with multiple symptoms and refusal to acknowledge a severe mental health problem are rare.

Table 1 Prevalence of Medically Unexplained Symptoms as the main reason for consulting a GP

Study	N	Location	Medically Unexplained Symptoms	Comments
Mumford(64) 1991	680 (attenders, any reason)	UK	5% probable MUS 10% possible MUS.	MUS more likely if past or current depression or anxiety
Peveler (17) 1997	170 (booked consultations)	UK	19%	10% had a mood disorder but presented physical symptoms only.30% had multiple somatic symptoms but only one third of these also had a psychiatric disorder
Melville(71) 1987	222 (new illness episode)	UK	Not specified at onset, 3% after 6 months	90% of physical symptoms, whether explained or unexplained by organic disease required no more than two consultations over 6 months
Palsson(72) 1988	78 (booked consultations)	Sweden	16%	8/13 with MUS met hypochondriasis criteria
Pilowsky(73) 1987	100 (booked consultations)	Australia	39%	Functional scored higher on scales of affective disturbance & disease conviction
Scicchitano(74) 1996	112 (New illness episode)	Australia	27%	Overall no difference between organic and functional in GHQ score. Male functional scored higher on affective disturbance and disease conviction (but N=5). No differences in females.

Table 2 Overlap of Somatisation and psychiatric disorders in primary care / community samples

	Sample size and type	Somatisation measure	Prevalence of conditions	Psychiatric morbidity in patients with unexplained symptoms	physical symptoms in patients with psychiatric illness
Posse(75)1998	406 consultants, Sweden	Somatisation scale	15% Somatisation on screening	25% Major depression 40% dysthymia 20% no psychiatric diagnosis	
Escobar(66)1998	1546 consultants, USA	Symptom checklist	22% had $\geq 4/37$ symptoms (men) or $\geq 6/41$ symptoms (women)	38% depression 8% dysthymia 10% generalised anxiety 36% no psychiatric diagnosis	36% had at least one psychiatric illness. 13% had psychiatric illness + abridged Somatisation 23% no Somatisation
Kisely(68)1997	5447 consultants in 14 countries	Reported symptoms categorised as explained or unexplained	6% 4-6 explained symptoms 13% 4-6 unexplained symptoms 2% 4-6 both	Rates of confirmed psychiatric illness: ≤ 4 explained or unexplained symptoms =10% >4 explained symptoms =33% >4 unexplained = 69% >4 explained and unexplained = 68% >10 unexplained or 12 explained =>80%	52% of those with psychiatric disorder had either >4 explained or >4 unexplained symptoms. Social disability worse if unexplained symptoms, physical disability worse if both explained and unexplained. GP detection of psychiatric illness higher if symptoms unexplained than if none or explained.
Toft(76)2005	1785 consultants, Denmark	Symptom checklist, diagnostic interview	36% somatoform disorder	OR for comorbid depression vs. non-somatoform = 1.5, OR for comorbid anxiety = 2.6	
De Waal(77)2004	1046 consultants + cross section	Symptom checklist, HAD scale	16% somatoform disorder	OR for comorbid anxiety and/or depression 3.3 vs. non somatoform	
Aggarwal(78)2006	2299 practice population	Self report of functional somatic syndrome	27% reported one or more syndrome	OR for high health anxiety 3.3 vs. no reported syndrome	

Lobo(79) 1996	1550 consulters, Spain	symptom count and reason for consulting	27% psychiatric cases with diagnostic interview	9% of population had psych illness but presented with MUS (over half of diagnoses were mild anxiety / depression) 20% <4 symptoms 65% 4-12 15% >12
Kirmayer(67;8 0) 1991	685 consulters, Canada	Symptom count and reason for consulting	11% had major depression or anxiety (n=75). 26% had one or more of >=4-6 physical symptoms, hypocondriasis, or somatic presentation of depression	Of 75 patients with major depression / anxiety: 17% presented psychosocial problem 41% presented physical but volunteered psychosocial basis 23% presented physical but acknowledged possibility of psychosocial basis 12 % presented physical and refuted psychosocial basis
Munk- Jorgensen et al(81) 1997	424 consulters, Scandinavi a	"Ill defined symptoms or mental illness with physical symptoms"	32% had psychiatric illness. 17% met Somatisation criteria described	151 Patients with psychological illness: 31% secondary to physical illness 29% presented coincidental physical illness 32% presented somatic symptoms 9% presented purely psychological symptoms
Feder et al(82) 2000	172 consulters, USA	"Physical complaints in excess of what would be expected..."	24% identified as multiple MUS by physician. 12.5% had current major depression 19% any current anxiety disorder	9/21 patient with major depression and 14/33 patients with any anxiety disorder were rated by their physician as having multiple MUS. "many patients with multiple MUS acknowledged mental health and social problems"

Characterisation of specific MUS syndromes

Several studies have shown overlaps between the syndromes which are included within the broad category of MUS. Many patients with irritable bowel syndrome, for instance, meet the criteria for chronic pelvic pain or fibromyalgia and vice versa. Analysis of population data to identify meaningful classes or disorders have variously suggested eleven(83), five(84), and four(85) symptom clusters. While such clusters appear broadly to fit clinical patterns, there appear to be no differences between them in terms of psychological characteristics and indeed there are many similarities: particularly common aetiological factors and responses to treatment. From this perspective, Deary(86) and Wessely(20) have argued strongly that individual symptoms, while connected to recognised syndrome clusters, are more strongly associated with a single unifying factor, possibly related in some way to the personality trait of neuroticism. Such a three level relationship is shown in Figure 1.

Aetiological Factors

A number of studies have attempted to identify specific aetiological factors for MUS although in general the aetiological factors for MUS are similar to those for anxiety and depression. Deprivation and childhood or family illness(87;88) may all play a part as may concurrent stress(89). In women with MUS there is a higher incidence of past or recent abuse(90), particularly in the case chronic pelvic pain in which around a third of patients will have some history of abuse. A longitudinal study of patients at age 36 and 42 years showed that physical symptoms at the first assessment predicted later mental health problems, and also that mental health problems independently predicted future physical symptoms(91).

Psychological processes in patients with MUS

While it seems clear from simple clinical observation that psychological factors are important, defining these has been more difficult and remains incomplete. Much has been made of the difference between patients with psychiatric illness who present somatically and those presenting psychosocially(56;87). Initial work was grounded in the belief that somatisation represented a flawed process in which failure to recognise the true problem led to ongoing distress for the patient (and high healthcare costs and frustration for the physician). A systematic review of eight studies of such patients in primary care failed to identify consistent differences between people with

psychiatric disorder who present psychologically and those who present with physical symptoms(92), except that generally those with physical symptoms were less distressed. However a separate study suggested that the level of psychological distress may be more to do with the number of physical and mental symptoms and unrelated to whether the individual acknowledged a psychiatric component to their illness(88). Even measures of cognition such as health anxiety or bodily awareness appear little different between “psychologising” and somatising patients with anxiety and depression. While only around a quarter of affected patients present with purely psychological symptoms, most of the remainder will accept the possibility of a psycho-social component to their physical symptoms even if they do not volunteer it within the consultation(93). Practical reasons (such as lack of time, or a sense that problems are not relevant or amenable to treatment) seem more important than failure to recognise their own mental distress in explaining why patients choose not to disclose psychological problems in consultations(94). Although major depression is harder to recognise if presented somatically(95), there is conflicting evidence that improving detection improves outcome.

As psychiatric illness labels, while common, do not apply to many patients with MUS, a variety of psychological characteristics and processes have been suggested. While many carry theoretical appeal, some are too broad, (the personality dimension of neuroticism(86)) or too restricted (alexithymia – the inability to express emotions in words)(96), to be useful in a heterogeneous population. This section examines four processes: hypochondriasis, somato-sensory awareness, attribution (including illness beliefs), and reassurance. To some extent these concepts overlap and patients who demonstrate one tend to demonstrate others. Nonetheless they may represent different facets of the problem and warrant further examination.

These processes appear to affect the decision to seek medical attention for any problem whether “organic”, such as a respiratory infection, or “functional”. Little and colleagues recently highlighted the importance of medically unexplained symptoms in influencing decisions to consult for all conditions(12). Studies of these processes which compare consultations with and without physical disease may underestimate their importance.

Hypochondriasis

Hypochondriasis is a preoccupation with fears of having, or the idea that one has, a serious disease based on misinterpretation of bodily symptoms despite appropriate medical evaluation and reassurance. It overlaps with somatisation but appears not to be identical: in a study of 184 primary care patients(97) 20% met criteria for hypochondriasis of whom two thirds also met somatisation criteria based on number of symptoms; a further 20% of the sample met somatisation criteria without hypochondriasis. In another study(98), hypochondriacal patients were more likely to interpret physical symptoms as being due to illness than patients with non-hypochondriacal anxiety, and in two separate studies of healthcare usage hypochondriasis was a predictor of repeated consultation, particularly in men(73;96). Robbins and Kirmayer demonstrated hypochondriasis in 10% of over 500 primary care consulters, about half of whom continued to show hypochondriacal beliefs a year later(99). Improvement in illness worry were matched by improvement in overall wellbeing, whereas persistence or new occurrence of hypochondriasis was most strongly associated with affective disorder. Hypochondriasis is a common feature of patients referred to secondary care with MUS(100) and also indicates a greater likelihood of symptoms persisting at follow up(101).

Somato-sensory awareness

Individuals have varying degrees of bodily awareness. The tendency to notice, and also to amplify, benign sensations is an element of patients with MUS. For instance in patients with palpitations but normal investigations(102) high levels of somato-sensory awareness predict persistence of symptoms. The cognitive model of panic disorder, which frequently coexists with MUS, includes awareness of bodily sensations which are amplified by the resultant anxiety(103) – for example, awareness of heartbeat or breathing triggers arousal which in turn increases heartbeat or breathing and sets up a cycle. Heightened bodily sensitivity is a feature in many patients with irritable bowel syndrome, fibromyalgia and chronic fatigue syndrome.

Attribution and Illness Beliefs

Attribution is the cognitive process whereby somatic sensations are interpreted in the context of the body and its physical and social environment. Using the example of fatigue, attributions can either be normalising (“I’m tired because I’m overworking

and unfit”) somatic (“I’m tired because my muscles have been weakened by a virus”) or psychological (“I’m tired because I have depression”).

Two small studies of frequent attenders in primary care(104) and patients with high health anxiety(99) suggest that difficulty in finding or applying normalising attributions may be important. Both groups studied were able to list fewer normalising explanations than healthy controls, and while the high health anxiety group did identify more possible illness explanations than controls, the frequent attenders did not. This difficulty in seeing symptoms as benign may explain why reassurance which rules out problems but does not offer alternative tangible explanations so often fails(105).

One of the few longitudinal studies of changes in health related cognitions(89) identified a pattern whereby symptoms occurring at a time of newly increased stress tended to be attributed to the stress. Only if the stress persisted did symptoms begin to be presented to doctors as possibly physical. This is compatible with the idea of patients being able to tolerate and normalise symptoms for a limited time before seeking assessment and reassurance that their original attribution was correct.

While doctors have medical models of illnesses, patients also have complex and broadly consistent lay models of health and disease(34). Consistent features of these include the name of the condition and its symptoms, the personal consequences of it, how long it will last and the extent to which it can be controlled or cured(106).

Patients appear to have health beliefs about individual symptoms as well as established diseases and Salmon(107) proposed eight dimensions: four covering aetiology (stress, environment, lifestyle and weak constitution), three concerning mechanism (wearing out, internal structure and internal function) and a final dimension of concern raised by the symptom.

Not only do patients have clear views about their symptoms in their own right, they also view their own experience of the symptoms as at least as important as a doctor’s opinion about them. Salmon and colleagues have demonstrated that patients perceive doctors as denying the validity of their symptoms(58) but that where doctors develop tangible and non-blaming models of conditions with their patients and form

constructive alliances against the illness, patients are then able to accept medical opinion(40).

While MUS tend to change over time(108), attributional style appears to be much more consistent(109). Changing specific attributions about symptoms appears to be important in effecting improvement(110).

Catastrophisation

Catastrophisation is a phenomenon which has been studied in chronic pain and MUS and represents an exaggerated negative interpretation of both actual and anticipated pain(111). It is seen as an over-reaction of natural threat avoidance procedures and is both influenced by mood and mediates some of its effects(112). It has neurophysiological correlates(113) and is associated with level of disability(114) and pain(115). Encouragingly it appears amenable to cognitive behavioural interventions and treatment is associated with improved outcomes(116).

Reassurance

Illness belief models explore how patients see illness as threatening. Doctors, through treatment and reassurance seek to reduce that threat. Unfortunately reassurance is not always effective: between a third and half of patients report continuing concern about serious illness after normal cardiac ultrasound or angiography(117). The effectiveness of reassurance appears to be related to patient characteristics. While all patients who received a normal result after upper gastrointestinal endoscopy(36) experienced immediate reassurance, those with the highest levels of health anxiety had returned to original levels of concern within one week and this persisted for a year. Similar transient effects have been observed for brain scanning in headache(37) and colonoscopy in irritable bowel syndrome(38)

Psychological models of threat reduction suggest two separate processes(105): one emotional-heuristic (calming, protecting and threat avoiding) and the other cognitive-systematic (information seeking and threat analysing). While emotional, threat avoiding, reassurance (which may be non-verbal as well as verbal) may be effective in alleviating distress in the short term, it may do nothing to weaken illness representations. If symptoms keep recurring, repeated use of this type of reassurance is likely to produce a cycle of reassurance seeking and giving that is self

perpetuating. (118) In contrast the cognitive model of threat analysing is more threatening in the short term, but more likely to produce long term changes which in turn can be associated with improvement(110). Research into minor physical illness suggests that patterns of doctor-patient interaction tend to be self-reinforcing(15) and that doctor behaviour in one consultation affects future consultations.

Treatment

Until recently there have been few studies of treatment of MUS in primary care and these used a variety of treatment methods including individual and group psychological therapies. Recently several groups have reported at least preliminary results of studies of broadly similar reattribution therapy delivered by GPs to their usual patients after additional postgraduate training.

Ten years ago, Morris and colleagues(119) devised a training package to help GPs recognise depression in patients with MUS and treat it. The outcomes of a before and after training comparison(119) suggested that patients who acknowledged their depression when it was pointed out to them showed improvements in depression and global function and there was a net reduction in healthcare costs(120). Agreement between doctor and patient predicted a good outcome while the patients who denied the possibility of depression did not improve, and felt that their doctors understood them less well after the intervention(121).

Lidbeck(122) and colleagues evaluated a programme of group cognitive behaviour therapy (CBT) after thorough physical examination, for patients with MUS in primary care. Thirty two subjects were contrasted with 17 waiting list (8 months) controls. At 6 months follow up there were significant changes in illness worry, illness behaviour, and medication usage in the early treatment group but no change in mood or social problems. No data are presented on subsequent consultation rates. An American study(123) of a rigidly structured behavioural intervention for patients with MUS involving 6 weekly sessions with homework demonstrated significant improvements in mood and physical symptoms both one week and six months after the course compared to waiting list controls.

There have been no studies in primary care of individual CBT although several secondary care studies have evaluated cognitive behavioural therapy and found it

effective in changing cognitions, wellbeing, and health service usage both for specific syndromes(110;124-126) and for an unselected population(127). Improvement in symptoms with CBT is not always matched by improvement in emotional state(52).

Other trials of psychological therapies have generally been small, however a large study of psychotherapy for patients with irritable bowel syndrome(128) showed sustained improvement in symptoms and wellbeing. A recent study of the disclosure of emotionally important events had no effect on patient's health(129).

A meta-analysis of antidepressant treatment for MUS(130) demonstrated beneficial effects in a wide range of conditions, although not chronic fatigue syndrome, with a number needed to treat of 3. Benefits were seen equally in those with or without depression. There was insufficient evidence to make detailed recommendations on optimal drugs, doses or duration of treatment.

Trials of reattribution

Several groups have now studied one form or another of the concept of reattribution: this was originally proposed by Goldberg(131) et al and more recently developed by Gask and Morris(119) and more recently Fink(46). Broadly speaking reattribution is a way of directing consultations towards linking physical symptoms and experience to psychological factors as a way of treating somatisation. As such it is based on the belief that the patient would be better off considering the psychological causes of their distress rather than seeking a medical explanation.

Benefits to date have been modest: Gask and Morris carried out a before and after study with 8 GPs and showed short term reduced healthcare costs(120), but patient wellbeing increased only in patients who acknowledged at the outset that psychological factors may have a part in their symptoms(119). Blankenstein(132) demonstrated the feasibility of teaching Dutch GPs to carry out reattribution with their own patients over 3 sessions, but have published limited outcome data.

Larisch(133) added reattribution to an existing psychosocial primary care programme in a German study with good short term results, but benefits were minimal at 6 months. Within that study, it appeared that contradictory effects became apparent however, one group of patients who recognised mental health difficulties at the onset

felt better supported by their GPs, but did not improve on the study instruments, while another group had low levels of insight at the outset and improved, but felt less comfortable with their GP(134). Rosendal(135) have published results on GP attitudes and confidence toward patients with MUS following an intensive extended reattribution model. This showed benefit for process measures of GP care but no patient outcome data are yet available. The MUST trial from Morris and colleagues(136) has demonstrated that reattribution can be taught in a shorter package than previously used, even with GPs not already experienced in applying a strong biopsychosocial model of care. Again, outcome measures are awaited. Recently a study of reattribution with and without psychiatric consultation and liaison(137) suggested a substantial benefit of the combination over reattribution and it may be that the two offer complementary treatment elements.

The importance of explanations

The importance of a good doctor patient relationship and of acknowledging patients' concerns has been demonstrated(39). Although there is no direct evidence of the effect of consultation behaviour on patients with MUS, the evidence that doctor behaviour for minor physical illness affects future consultation rate(15), and that a positive, patient centred, approach(138) improves satisfaction and enablement, and reduces symptom burden and health service usage strongly point to this being important. In qualitative studies of patients with MUS, Salmon(40) identified three types of medical explanation: rejecting (in which the patients perceived the doctor as denying the reality of their symptoms and in which there was unresolved conflict over explanations), collusive (in which the doctor gave in to the patient's interpretation of symptoms but in doing so lost the respect or trust of the patient) and empowering (in which the doctor provided tangible, non-judgemental, explanations which legitimised the patient's suffering and offered opportunities for self management). The empowering explanations were distinctive in that patients regarded them as valuable foundations on which to build recovery, or at least cope with their condition in partnership with their doctors.

Conclusions

The notion that most medically unexplained symptoms are the result of a single process of somatisation (particularly the somatisation of mental distress), or are due to a somatisation disorder which can be defined primarily in terms of numbers of symptoms, is no longer supported by the evidence. There is now good evidence that physiology, personality, life experiences, health cognitions and interaction with healthcare professionals are all important and any new paradigm needs to include them.

Chapter 3 Electronic diaries for symptoms research

Introduction

The research in this thesis is based on data from electronic diaries. This chapter describes a systematic literature review of the use of electronic diaries in symptoms research. It addresses issues of data accuracy, user acceptability, and the capacity of diary study designs to answer important questions in symptoms research.

Diary studies, in which patients record their symptoms or thoughts regularly, have been a feature of psychosomatic research for over 25 years(139). Traditionally they have been completed on paper and have been used to track the variation of a single symptom or the interactions between symptoms and other variables such as mood(140) or stress(49).

Compared to conventional questionnaire studies, diaries are less prone to recall bias(141). Recall bias is potentially a major confounder in symptoms research, as retrospective recall is affected by both current state and experiences since the recalled event(32). However, diary studies are associated with their own unique problems and there have been doubts about both the accuracy and timing of responses, particularly their susceptibility to being completed from memory after the stated time(49;142).

Electronic Diaries

Since the early 1990's developments in handheld computer technology have offered an opportunity to exploit electronic diaries, on which users enter data via a touch-screen and stylus. User-friendliness is achieved by selecting options from on-screen lists or completing Likert-type or visual analogue scales (VAS). All entries are date and time stamped and data is stored in the device for later retrieval either by connection to another computer for download or by telephone. Electronic recording of symptom information has higher start-up costs than other methods (handheld devices typically range in price from 100-400 Euros or US dollars and require additional modification and programming before use). However a number of free and

commercial software programs are available for developing handheld computer diaries(143) and good principles of electronic diary design have been published(144).

Strategies for collecting data

Traditionally, paper diaries have been completed once daily, typically toward the end of the day, but this procedure is prone to recall bias, with selective memory for both the most severe and the most recent events(141). In recent years Stone and colleagues have pioneered a technique known as ecological momentary assessment(145) in which multiple data entries are requested each day and the subject is asked to record how they feel at that point in time. Such studies have been particularly valuable in addiction research where moment to moment changes may have major implications for relapse(146). To avoid repetition at set times of day, electronic diaries for this type of study are set to sound an alarm at semi-random intervals.

Methods of analysis and reporting

Electronic diary studies generate rich data with many repeated measures of multiple variables, typically with 50-150 time points. While simple monitoring of trends may be sufficient for descriptive studies of day to day variation, more complex study designs, such as those seeking associations or causal sequences require particular caution in interpretation. Most time series data from diary studies show marked autocorrelation(141;145) whereby values depend on the preceding values in the series. This violates the assumption of independence which underlies parametric statistical methods such as multivariate regression(31).

Three main analytical approaches are used. In the first, the data are treated to remove autocorrelation and trend, either by differencing or use of specific time series modelling techniques such as autoregressive moving average (ARMA) models(49). The second approach uses multilevel modelling techniques in order to carry out the analysis at both within-person and between person levels(31;147). Finally there are specific true multivariate time series analysis methods which require both long data series and considerable statistical skill(148). New techniques from non-linear science have potential but have not yet been adequately tested(149).

Systematic Review Aims

I conducted a systematic review of electronic symptom diary studies in pain and symptoms research which either evaluated their accuracy and usability or investigated the associations between physical and psychological variables. The aims were (a) to determine whether electronic diaries generate valid data for symptoms research; and (b) summarise the evidence regarding the interaction of somatic and psychological symptoms that has been acquired from electronic diary studies.

Methods

Searches were carried out of Medline, Embase and PsycInfo databases for studies of symptom disorders using electronic diaries between 1985 and November 2005. The specific search criteria were for intersections of the following: (pain or symptom\$ or somatoform) and (diary or diaries or handheld) and (computer\$ or electronic).

Relevant additional studies identified from references were reviewed. A further search was made for electronic diary studies in psychology and using the term “ecological momentary assessment”. Studies using only pen and paper diaries were not systematically reviewed although their findings are relevant to the way in which electronic diary data is handled. Studies were excluded if they did not include physical symptoms (for example diary studies in addiction research) or if the electronic diary was simply used to repeatedly measure symptoms within a drug trial

From the initial screening process, studies were divided into two groups. Group 1 comprised those electronic diary studies which tested associations between one or more physical symptoms and psychosocial variables, either in normal life or during a therapy programme. Group 2 comprised those studies which assessed electronic diaries solely in terms of acceptability, accuracy or compared to pen and paper diaries.

Information about acceptability, accuracy and comparability of electronic diary data was extracted from both groups and used to address the first aim of the study regarding the validity of data from diary studies. The second study aim, summarising the interactions between symptoms and other variables obtained from diary studies was addressed by reviewing studies in group 1 only and inductively developing a

descriptive classification. No formal meta-analysis of results was carried out as few studies were directly comparable.

Results

The search strategy identified 32 papers from 24 studies. It also identified three narrative reviews of methodological issues in diary studies(141;150;151) and a set of reporting recommendations(31). Studies were allocated into one of two groups based on whether they reported interactions of symptoms and other variables (group 1), or whether they were concerned solely with the accuracy or usability of electronic diaries (group 2). Twenty one papers from 15 studies were allocated to group 1 and are summarised in table 1. The remaining 11 papers, allocated to group 2, are summarised in table 2.

Conditions or symptoms studied

The 15 studies in the group 1 related to a range of conditions: chronic pain (5); fibromyalgia (4) ; temporomandibular joint dysfunction (2); migraine (2) ;multiple chemical sensitivity(1) and asthma(1). The 9 studies in group 2 related to pain (5), overactive bladder, menstrual symptoms, heartburn and mood (one each).

Description of studies and diary devices

The devices used for electronic diaries varied. Some earlier studies used Psion organisers with a keyboard and rectangular screen presenting a few lines of text, but most studies used a touch-screen based device such as a Palm™ handheld computer. These had a graphical display and required input by touching the screen with a stylus. All devices shared the ability to sound an audible reminder when entry was due and to date and time stamp all entries. Data were stored electronically within the device and in most studies extracted at intervals between 1 and 3 weeks at a meeting with the researcher. Diaries used a range on-screen methods of data collection and these are listed in Table 3 and Table 4 .

Sampling intervals and study duration varied considerably. In general, shorter studies used more intensive monitoring (six studies used between 5 and 8 daily prompts) while longer studies tended to use once or twice daily recording. Most sampling strategies were pre-arranged: either at fixed times or in semi-random fashion, with the diary program introducing an element of unpredictability in the timing while still

ensuring a reasonable spread through the day. Three studies used event based sampling during peaks of symptoms(152-154) in addition to more structured sampling.

The number of items of data requested at each entry varied between a single item in some feasibility studies, to a complex profile of pain and other variables.

In studies seeking associations between variables a number of analytical techniques were used including cross correlation of variables over time(155), regression after adjusting for autocorrelation(156), multilevel modelling (115;157) and multivariate time series analysis(148).

Table 3 Electronic diary studies of associations between physical symptoms and psychosocial variables

Study	Context	Diary data collected	Data Input	N	Female (%)	Duration (days)	Intensity (entries / day)	Adherence	Timing	Analysis
Aaron et al 2004(115; 158-160)	TMD patients pain clinic; results of RCT of brief CBT vs. self care management	Pain, pain related beliefs and coping, catastrophisation, and mood, beliefs, activity	numeric 0-10 with end verbal anchors on touchscreen device.	71-126*	86	14-56*	3	98%	Fixed	MLM within and between persons over time. Papers includes short studies (1-3 weeks) on persons waiting for start of trial therapy
Feiler et al 2005(148)	Research FM clinic	Pain intensity, mood, sleep, self efficacy	11 point VAS on touchscreen device	43	100	84	1 (evening)	NS, (8 subjects rejected)	Self	Multivariate time series technique: correlation of the frequency spectra to seek similarities between series Pooled data between patients for analysis
Saito et al 2005(154)	MCS specialist clinic (+ normal controls)	Specific Symptoms + Psychosocial Simultaneous air sampling for chemicals	Adjustable 21 point scrolling bar length on wristwatch for computer	14 cases 12 controls	50	7	2 + symptoms	100% (1 subject rejected)	Random + at time of symptoms	MLM grouped by patient and whether experiencing symptoms.
Liszka-Hackzell & Martin 2004(155)	patients with back pain for >6months or <2 wk attending hosp clinic	pain, activity by actimeter	numbered key 0-10 on key based data recorder	30 (15+15)	40	14-21	10	"complete" (additional 11 pts rejected)	fixed 90 mins + exp	Activity (recorded per minute) and self reported pain (every 90 minutes) smoothed to give equivalent 10 minute samples, cross correlation with lags.

Stone et al. 2004(157; 161;162)	Research Unit; inflamm & non-inflamm musculoskeletal pain	Pain severity on several dimension; current location / activity	100 item VAS; word checklists branching logic depending on initial question on touchscreen device	66 91 in study of compliance	85	14	3-12, mean 7)	94	Random	MLM.
Litt et al. 2004.(163)	TMD service advert	Pain, control, coping, catastrophising, mood	11 point likert scales on scrolling key based device	30	87	7	4	81	Random	MLM within and between person, with adjustment for autocorrelation to identify prior, concurrent and later associations.
Roelofs et al. 2004(164)	pain clinic	Pain, pain related fear, attention to pain, fear of movement.	7 point Likert scales with verbal anchors, on touchscreen device	40	57	7	8	76	Random	MLM within and between person.
Giffin 2003(152)	Migraine: clinic	Migraine, related symptoms	VAS for current severity, checklists of symptoms, headache related behaviour on touchscreen device	88	95	90	1	>80 (5 excluded)	Random + at time of symptoms or headache	Probability of future migraine after report of premonitory symptoms.
Affleck et al. 1996,(156)	FM: rheumatology clinic	Pain (multiple areas) & attention to pain, sleep quality	7 point likert with verbal anchors on scrolling key device	50	100	30	4	99	Random	MLM within and between person.
Zautra, Affleck et al. 2001(165; 166)	FM: rheumatology clinic + media	Pain (multiple areas), fatigue, 16 item mood scale, daily goals record	7 point likert with verbal anchors on scrolling key device	89	100	30	3	99	Random	MLM

Peters 2000(167)	pain clinic, equal subgroups <12 or >12 months pain	Pain, physical function, pain related cognition, social support	7 point likert with verbal end anchors on touchscreen device	80	78	28	4	83	Random	MLM within and between persons .
Honkoop 1999(153)	Migraine without aura: headache clinic	headache and associated symptoms	Unspecified methods on touchscreen device	56	100	70	6	80	Random	Descriptive analysis of experience of migraine episodes.
Viane et al 2004(168)	Chronic pain clinic / FM support group	Pain, attention to pain, goal directed activities, baseline acceptance of pain	7 point likert with verbal end anchors on touchscreen device	62	NS	14	8	88%	Random	Averaged diary data over time for between person only analysis
Affleck et al 2000(169)	Asthma, university hospital clinic	Mood asthma symptoms, peak expiratory flow	11 and 7 point likert scales on scrolling key based device	48	65	21	3	98%	Random	MLM within and between persons.
Kop et al 2004(170)	FM and / or CFS recruited via media	Pain, fatigue, distress, objective measure physical activity (actigraph)	11 point rating on wrist-worn actigraph keypad	38	74	5	5	NS	Fixed	Within subject estimation of cross correlation between exercise and prior / subsequent symptoms.

Table 4 Studies evaluating electronic diaries solely in terms of acceptability accuracy or compared to pen and paper diaries

Study	Context	Diary data collected	Data Input	N	Female (%)	Duration (days)	Intensity (entries / day)	Adherence	Timing	Analysis
Gaertner et al 2004(171)	Chronic Pain	Pain severity	11 point numeric rating scale on touchscreen device	24	NS	14	1	92%	Fixed	Crossover study with paper diary 83% preferred electronic diary
Williams et al 2004(172)	Fibromyalgia	Pain severity	21 Point verbal descriptor on touchscreen device	14	93	84	5	85%	Random	Feasibility study of frequent sampling over long period of time.
Lauritsen et al 2004(173)	Dyspepsia	4 symptoms	4 point verbal descriptors on touchscreen device	54	52	28	2	82%	Fixed	Comparison with paper and phone diaries (Randomised, similar numbers other arms). Higher satisfaction than phone diary.
Kreindler et al 2003(174)	Mood	18 mood related items	VAS with verbal end anchors on touchscreen device	28	66	NA	NA	NA	NA	Comparable to 10cm paper VAS for accuracy and transposition errors. Formal test-retest showed Cronbach's alpha = 0.89
Quinn et al 2003(175)	Patients with overactive bladder	Urinary symptoms, events, volume etc	Event based recording using	35	90	7	Event based	NA	Unspecified	Crossover with paper diary, 94% rated electronic diary as easy to use.
Stone et al 2003(142; 176)	Pain	Chronic pain	20 items, 7 point Likert scales	80	50	21	3	94%	Fixed	Case control comparison electronic diaries vs. paper with pressure sensor. 32% of days, reports completed on a different day. Only 11% of paper entries at stated time.
Jamison	Healthy volunteers	Experimental sensory stimuli	VAS with verbal end anchors only.	24	79	NA	NA	NA	NA	Psychophysical investigation; electronic VAS "remarkably similar to a VAS on paper"

et al 2002(177)	Low back pain	Pain severity now and over each of preceding 16 hours	on paper and on touchscreen device	24	45	365	1	71%	Fixed	Feasibility study of 1 year data collection. Daily entry of retrospective hourly pain on 71% of days.
Jamison et al 2001(178)			VAS with numerical score							
Johannes et al 2000(179)	Healthy volunteers	Menstrual symptoms	Choice of verbal descriptors	23	100	28	1	96%	Fixed	Feasibility study crossover: one month paper, one month electronic diary. Higher data entry rates and 70% preferred electronic
Lewis et al 1995(180)	Pain patients	Pain	Single measure on pushbutton device	36	58	70	4	NA (10 subjects withdrew)	Fixed	High correlation between paper VAS and pushbutton device rating at baseline. Diurnal variation in pain.

Validity of electronic diary data for symptoms research

Data on compliance, accuracy and acceptability of electronic diaries was extracted from studies in both groups. All but two studies reported compliance rates and these ranged from 76 to 100% of possible entries made. However these were inflated in some studies by excluding participants from analysis if less than a certain percentage of possible entries were completed. No studies reporting this excluded more than 25% of participants.

Several of the evaluation studies addressed the accuracy of data entry on electronic diaries. Although visual scales are constrained in length by the screen size (usually approximately 5x5cm in contrast to the standard VAS 10cm line) results matched those from the paper VAS closely in the two validation studies which compared these(174;177). Although these two studies examined test-retest reliability, few of the association studies reported this. In studies which measured acceptability or preference, few participants appeared unable to use electronic diaries and most preferred them to pencil and paper diaries(171;175;179).

The key study comparing accuracy of electronic and paper diaries comes from Stone and colleagues(142;142). They fitted a pen and paper diary with a hidden pressure sensitive electronic device which logged the actual time at which each entry was made. While participants' self-reports of completion (judged by the date and time they had stated for each data entry) were similar to those obtained automatically from electronic diaries, only 11% of pen recordings were actually made within 15 minutes of the time stated.

Electronic diaries did not generally show reactivity(158;161), although in one study the diary appeared to produce a lower overall burden of symptoms than compared to written recall on either a daily or weekly basis(162).

In summary, electronic diaries have been shown to generate valid data for symptoms research.

Evidence of from diary studies in symptoms research

Review of the studies in group 1 yielded a wide range of measures and interactions. These were grouped into five themes which are listed, with specific examples, in Table 5

Table 5 Key findings of diary studies

Category	Condition	Observation
Experience		
	Multiple chemical sensitivity	typical symptoms occurred only when air sampling revealed the presence of trigger chemicals(176)
	Migraine	Sufferers accurately predicted future attacks from non-specific prodromal symptoms(174)
	Migraine	Sampling confirmed historical assessments of duration(175)
Interaction		
	Pain	Pain varies with time of day(189). Positive and negative affect are differently related to pain(187;188). Low mood & pessimism have greater effect on relatively pain free days(179;188)
	Asthma	modest effect of mood on respiratory function in asthma(191)
Sequential		
	Chronic Back pain	No relationship found between exercise and pain(177) (in contrast to acute back pain within one week of onset)
	Fibromyalgia & CFS	No relationship found between exercise and subsequent pain(192)
	Fibromyalgia	poor sleep and low mood both predicted increased pain the next day(178)
	TM Dysfunction	pain was influenced by several variables currently and at the preceding entry(185)
Process		
	Back Pain	Catastrophisation was associated with current(189) and future pain(186)
	TM Pain	Catastrophisation was associated with pain(111;185)
	Chronic Pain	Pain related fear was a determinant of pain(186)
	Fibromyalgia	Self efficacy was associated with recovery(170)
	Chronic pain	higher reported activity and lower attention to pain were found in those showing acceptance of pain at study outset(190)
Intervention		
	TM Dysfunction	catastrophisation & perceived control over pain both improved more in patients treated with short course CBT than self help(181)
	Fibromyalgia	changes in self efficacy were central to improvement(170)

Experience effects were those where variables were related to specific experiences. Patients with multiple chemical sensitivity were shown to have typical symptoms

only when air sampling revealed the presence of trigger chemicals(154) and patients with migraine were shown to be able to predict future attacks from non-specific prodromal symptoms(152). Sampling of the hours after a migraine headache confirmed historical assessments of duration(153). In each study the diary method produced unique information by collecting information in a way which avoided recall bias, which, particularly in the prodromal migraine study would be inevitable with any other study design.

Context effects included assessment of the modest effect of mood on respiratory function in asthma(169), the contribution of mood to the experience of pain at different intensities(181) and the effect of personality traits such as positive and negative affect on pain(165;166). The studies of mood and personality suggest that low mood or pessimism has a greater effect on relatively pain free days(166;181). These studies provide justification for more complex models of the interaction between pain and affect than simple cause and effect. Several studies demonstrated that perceived severity of chronic pain increased through the day (167).

Sequential effects were shown by studies that analysed data as an ordered time series. Two studies addressed the issue of pain and exercise, in both cases by combining objective measures of activity in addition to electronic diaries(155;170). While patients with acute back pain (within the first week of treatment) showed a relationship between exercise and subsequent pain, this did not appear in patients with chronic back pain(155) or fibromyalgia and CFS(170). In patients with fibromyalgia, poor sleep and low mood both predicted increased pain the next day(156), while TM dysfunction pain appeared to be influenced by a wide range of variables both currently and at the preceding entry(163).

Both these two studies measured contemporaneous physical symptoms and psychosocial variables; both showed that the web of interactions was complex and did not fit a simple sequential model. While statistically significant associations were identified in both it is important to note that they were identified post hoc rather than specifically predicted from an earlier data set and prospectively tested for.

Electronic diary studies lent support to the importance of four separate cognitive processes. These were identified before the data collection by cross-sectional questionnaire and then matched against symptom burden (typically pain) recorded in the diary. Because of the difficulty in operationalising an established question instrument to simple repeat measures items these were not recorded in the diary. Catastrophisation was associated with current and future back pain(164) and TM pain(115). Pain related fear was shown to be an important determinant of pain(164) as was self efficacy in patients with fibromyalgia(148). Acceptance of pain at the onset of a diary study was associated with higher reported activity and with lower attention to pain(168).

Two of the process measures also feature in the two studies which used diaries to monitor cognitive behavioural interventions. Turner et al showed that catastrophisation and perceived control over pain both improved more in patients treated with short course CBT than self help(159), while Feiler found that, during therapy for fibromyalgia, changes in self efficacy were central to improvement(148). Feiler's study in particular used well thought out strategies for analysis using multivariate time series tests of cross correlation and cross spectral coherence to show associations, but not to demonstrate cause.

These studies appeared well planned and executed, with appropriate methods of analysis. Their findings however, while often achieving statistical significance, were generally modest in magnitude. Often relationships between variables appeared to be bi-directional: for instance with increased pain predicting lower mood but lower mood also predicting increased pain(163).

Discussion

This review indicates a growing literature that has examined or reported on studies using electronic diaries. It suggests that electronic diaries are both acceptable to users and sufficiently accurate for research purposes. Data from electronic diaries is easy to extract and, with care, is amenable to a variety of methods of analysis. Studies that have compared electronic with pen and paper diaries have raised sufficient doubt over the accuracy of the latter that results from these should be viewed with caution.

However, although measures to confirm accuracy such as repeating certain questions are not difficult to implement in electronic diary studies, they have been used relatively infrequently.

Issues still to be resolved with e-diaries

There are still a number of unanswered questions about electronic diaries as research tools.

There has been no published assessment of the optimal number of items of data at each entry (and its corollary, the time required to complete an entry).

There has been no work comparing visual and word based data entry schemes. Earlier work on pain severity suggested individuals are able to discriminate 10 to 20 different levels, although in practice electronic diaries using simple visual analogue scales can generate even higher definition data by recording the exact location of a mark on the screen.

While reactivity studies have shown no sign of consistent trend in recordings over time, no studies have reported on whether the variance of readings diminishes with time. It is possible, at least in theory, that repeated use, particularly of word and number based ratings, may lead to consolidation on a small number of points as the study progresses. Future studies should report whether variability changes over time. Similarly habituation to questions in the same order may have unwanted effects on data accuracy; despite the flexibility of programmable devices to randomise the order of questions no studies reported doing this.

The optimal amount of training and support for diary users has not been determined. Some studies, implement detailed tuition and explanation programmes, others much simpler schemes. A recent review of electronic diary design suggested a maximum training time of five minutes(144).

Further, a number of disadvantages in using electronic diaries have been identified. The first is that participants need to be confident, willing and able to use them. The cost is not insignificant, but can be set against the savings in transcription time for data, and the possibility of re-use of devices in future studies. The risk of device

failure should be considered in planning studies and measures put in place to regularly back up data, ensure that batteries are adequately charged, and ensure that any problems can be quickly resolved.

Future research questions for e-diary studies

The current challenge for electronic diary studies is to define research questions which are too complex to answer by simpler designs, but are sufficiently simple to give a meaningful answer from the data it is possible to collect. Such questions will involve processes and interactions which are stable over time periods of several hours and which change relatively consistently in response to events or other measures, yet are too variable to be reliably predicted from single completion questionnaires.

In the field of symptoms research, electronic diary studies are well suited to investigating links between physical symptoms and a host of variables including daily hassles, symptom related cognitions, and low mood. It is also possible that simply through the reflective act of recording data, patients will gain insight and benefit(182).

Conclusion

Electronic Diaries are a powerful and efficient tool, fit for the purpose of collecting complex self report data. By recording the marked variation of symptoms and related constructs they offer a unique insight into the processes of symptom experience.

These properties make electronic diaries an appropriate tool for carrying out a longitudinal study into the variation of symptoms and psychological states in patients with MUS.

Chapter 4 Introduction to the analysis of time series data.

This thesis tests hypotheses about the interactions of mood and physical symptoms over time. This chapter acts as a brief introduction to conventional time series analysis. It introduces the concepts of autocorrelation (the dependency of current values on prior ones) and ways of dealing with this. It also sets the scene for the more novel concepts of non-linear dynamic systems in Chapter 5.

The opportunity to collect repeated data from one individual offers several advantages to the researcher. In particular it provides an attractive way of collecting large quantities of experimental data from relatively few subjects and offers the opportunity for sampling in several different circumstances. With those advantages also come challenges, for sampling and for analysis, which require care and planning. This section introduces the basics of time series analysis, in particular of self report data, taking up from the data collection methods in the previous section. It deals with the special considerations needed for time series data; issues of trend, periodicity and autocorrelation; specific time series models; and finally the problems of multivariate time series and implying causality from time series models.

Specific details of methods will be omitted from this introduction and be outlined directly before the relevant results to facilitate clarity.

Special considerations for time series data

While time series data offers the advantage of large datasets, it brings the disadvantage of clustering, in that the data do not have the independence of a conventional dataset. Hence conventional statistical methods such as correlation, regression and analysis of variance are not appropriate unless adjustments are made to counter this loss of independence.

While full data collection is ideal for any form of analysis, it is particularly important for a time series where missing data, or irregular sampling intervals, may change the apparent behaviour of the data being studied. While a number of methods are available to fit missing data they inevitably involve assumptions about what those values might have been had they actually been recorded which depend on how stable the data series is, and in turn how well a time series model can be fitted to it.

Trend, periodicity and autocorrelation

Methods of analysis of time series which can incorporate the dependencies of the data depend on the series possessing the property of stationarity. A process is stationary if its statistical properties do not change over time: thus the mean and variance of subsections of the series will be the same and there will be no overall trend. This is sufficient for most analysis and is referred to as second order stationarity or weak stationarity. Clearly some time series will have a cyclical component, for instance a woman's menstrual cycle may affect several physical and psychosocial variables and this needs to be accounted for in specifying time series model if appropriate.

The dependence of one point in a time series on those that have preceded it is termed autocorrelation. Analysis of the time series can categorise the extent of autocorrelation, in particular the size of its influence and the number of preceding values which affect the current value. Analysis of autocorrelation is used in this thesis to estimate appropriate ways of achieving a reasonable degree of stationarity and independence of data points.

Another method which can be used to describe a time series is to model it as a moving average process, using one of a range of equations or filters at each point to smooth out the short term changes. Typically they calculate a weighted local average with greater weight given to the adjacent values.

Dealing with autocorrelation:

When data displays autocorrelation, it is important to account for it before carrying out further analysis.

The simplest method, where it is more important to generate usable data which is free of autocorrelation than it is to specify the model, is to take the differences between each pair of consecutive points and use them as a differenced series. This differenced series can be tested for autocorrelation to ensure it is appropriate to use in regression analysis.

Autocorrelation

A time series can be expressed mathematically as a series X , with a value at point t of X_t . Preceding values are referred to as "lagged" such that the lag n value is X_{t-n} .

Autocorrelation is the process by which a variable in a time series (X_t) is dependent to some extent on its lagged values ($X_{t-1}, X_{t-2}, \dots, X_{t-n}$).

The autocorrelation coefficient at lag 1 can be expressed as:

$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^{N-1} (x_t - \bar{x})^2}$$

and the autocorrelation coefficient at lag k as:

$$r_k = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N-k} (x_t - \bar{x})^2}$$

From these, the autocorrelation function (ac.f) can be derived as the series of autocorrelation coefficients at lags (1...k) for the time series.

An alternative way of measuring autocorrelation is to calculate the autoregression coefficient α for a given series. In the case of an first order autoregressive model with lag 1 (AR(1)), the series can be represented as :

$$X_t = \alpha_1 X_{t-1} + Z_t$$

where Z_t is the error term at time t . Typically, autoregression coefficients near 1 are associated with slow decay in the ac.f, while lower coefficients nearer zero are associated with autocorrelation over very few lags.

Modelling time series, ARMA models

Going a step beyond simply measuring autocorrelation it is possible to generate a more accurate model of the time series. Using knowledge of long term trend, and a combination of moving average and autocorrelation it is possible to model a best fit curve to a time series. This is done by first removing any linear trend, then using an algorithm such as the Box-Jenkins ARMA (AutoRegressive, Moving Average) process. The result is a best fit model and a set of residuals should have sufficient independence to use in regression analysis against other variables.

This method is computationally complex, although readily available in statistical software, and requires testing to find the optimal solution for an individual time series.

Multilevel Modelling and autocorrelation

Many health-related researchers now use hierarchical or multilevel models of analysis of complicated datasets in which a series of nested analyses are carried out at different levels, for example associations may be tested at the levels of time (e.g. week of study), participant, and treatment condition.

While originally designed for analysing cross sectional data, multilevel models are still potentially valuable for analysis of diary data, however it is important that autocorrelation is taken into account within the models (183). While this can possibly be done by stipulating an inclusive model (e.g. by day, time, participant and group)(167) it is probably better done by including formal autocorrelation techniques.

Analysing time series in the frequency domain

A time series is generally thought of as a series of discrete points in time, it can, however be represented differently, as a curve described as the pattern produced by a series of interacting waves of different frequency and amplitude. This analysis in the frequency domain builds a model of the time series analogous to the overlapping sound waves of different pitch which represent a musical note, with resonances and overtones.

By decomposition of the pattern into component frequencies, a process carried out mathematically by the fourier transformation method, a frequency spectrum can be generated and analysed.

This spectrum can be used to compare time series which share similar patterns using methods such as cross spectral coherence(148).

The problems of multivariate time series

Having isolated the part of each point which is attributable to the underlying process, whether by applying a sophisticated ARMA model or simply by differencing, it is then possible to address any statistical associations with the residuals representing

the difference between the actual data points and the fitted values. Provided these residuals show no autocorrelation, they can be used in conventional statistical tests of association.

While it is reasonable to test for correlations between these residuals and other data, it is more difficult to fit a multivariate time series model in which data from two or more variables contribute not just to their own value over time, through autocorrelation, but to each other's, through cross correlation. The issue of independence of data points becomes extremely complicated when two data series become mutually dependent.

Implying causality from time series data

The problem of mutual interdependence in multivariate time series adds to the general difficulty in imputing cause from a statistical association.

A different approach to assessing sequential interactions in interrupted time series data was developed by the Nobel Prize winning economist Granger(184). He proposed that causality could be implied if the unexplained variance of one variable could be reduced by knowledge of the lagged (preceding) values of a second. He further argued that when this applied to a pair of variables in one direction only (knowledge of earlier A improves prediction of B but not vice versa) then this implied directional causality; the relationship can be described as A granger-causes B. In practice however granger causality is often bidirectional.

This can be expressed as a significant difference in error between the two regression models using preceding (lagged) values to predict the current one

Model 1(lagged Y):

$$Y = \beta_1 \text{Lag}(1)Y + \dots + \beta_n \text{Lag}(n)Y + E1$$

and Model 2 (lagged Y + lagged X):

$$Y = \beta_{y1} \text{Lag}(1)Y + \dots + \beta_{yn} \text{Lag}(n)Y + \beta_{x1} \text{Lag}(1)X + \dots + \beta_{xn} \text{Lag}(n)X + E2$$

Where $\beta_{y1\dots n}$ and $\beta_{x1\dots n}$ are regression coefficients for lags 1 to n of two time series Y and X respectively and E represents the unexplained variance. The difference

between these two models is expressed as an F–statistic with probability p from the comparison of the two models by the Wald test.

Granger causality has been applied to some biomedical time series(185), typically from EEG recordings with limited effectiveness(186). Currently non-linear methods of implying causality such as transfer entropy(187) are being investigated with complex biological signals, but are a relatively immature technique with little evidence of value in relatively short (<1000 points) time series.

Granger causality has not previously been previously investigated in the context of symptoms research.

Comparing individual time series

It is usual in time series analysis to develop one model for each series, selecting the best fit between data and model. When several series of similar data are generated there are a number of options for analysis.

One approach is to use multilevel modelling, with some form of correction for autocorrelation as discussed, as a way of combining individual series into one larger pool of data from which statistical effects are tested through nested analysis. In effect this treats each participant as one cluster of data, in the same way that one might treat one time period as a cluster within each participant or a group of participants with a shared characteristic such as a week as a cluster within the whole population.

An alternative is to combine individual series into one larger series, as was done, largely on grounds of obtaining an adequate sample size rather than from methodological choice by Feiler and colleagues(148).

Finally analysis can be carried out at the individual participant with the results then entered into a meta-analysis. This approach, of pooling separate idiographic studies, is not commonly used but is both theoretically justifiable and technically feasible(188). One potential problem when used with ARMA models is that it is possible that the best fitting model for one participant may not be the same for another, and comparison is easier if models match. Calculations based on simple differenced data, while not providing as highly specified model, have the advantage of all having been derived in the same way.

Conclusion

Time series data, while valuable, require specific forms of caution in their analysis. The dependence on values at one time on those on another means that conventional statistical tools such as regression can only be used when that dependence has been removed. Techniques such as ARMA are valuable for identifying underlying patterns and trends, but also have requirements for a degree of stability, or stationarity, of the data.

This section does not discuss specific examples of time series methods in symptoms research, rather these are discussed in Chapter 20.

Chapter 5 Complex systems and non-linear dynamic models in medicine.

This chapter introduces a novel approach to understanding and modelling complex phenomena using complexity science and the related study of non-linear dynamic systems. This approach, which shares some of its origins with the rather popularised “chaos theory” has been shown to be useful in explaining physiological systems, and this chapter develops its ideas as a way of modelling MUS.

In outlining the biopsychosocial model almost 30 years ago(189), Engel drew on the then recent recognition in biology that the behaviour of cells, organisms and ecosystems could be conceptualised in terms of an integrated systems theory, with interacting processes and contexts, rather than through a more conventional reductionist approach. With the development of computers capable of demonstrating both that complex pseudo-periodic chaotic behaviour of systems could be generated by very simple mathematical equations(190) and that simulations of multiple “agents” following rules for behaviour which depended on the state of their neighbours(191) could generate complex and responsive emergent behaviour , researchers have sought to address problems in biology and healthcare using methods from the science of chaos and complexity(48;192).

Models of complex and chaotic systems both belong to a broader family of theoretical models called non-linear dynamic systems. The elements of the name imply firstly that changes in the system are not linear (that there is no direct and proportional association between the level, or change in level, of one variable and another); secondly that the system is dynamic or persistently changing; and thirdly that the system has the other necessary properties of a system: a degree of coherence of the parts, separation from the environment, and recognisable inputs, outputs and behaviour.

The term non-linear dynamic systems covers a range of models, from the relatively simple to the immeasurably complex. At the relatively simple end of the spectrum are some forms of mathematical chaos(190): a pattern of non-periodic system behaviour which, in the pattern of its changes, appears both random and yet at times almost regular. Some of these systems can be generated by the reiteration of

remarkably simple mathematical equations. While researchers have sought such simply determined phenomena in biology and medicine, they have not generally found them. However several characteristics of the simplest deterministic chaotic systems do consistently appear in the complex systems of the real world as well as in mathematical models. Five of these properties are described in Table 6.

Table 6 Characteristics of non-linear dynamic systems

Phenomenon	Description	Example
Scale – independence (also called fractal scaling)	Patterns appear similar at different scales	The branching structure of bronchi is remarkably similar at all levels between main bronchi and terminal bronchioles(193) ¹ ;
Increasing divergence	If the system is mapped as a trajectory of points in one or more dimensions, two sections of the trajectory which initially overlap will diverge increasingly over time	(Very) short term prediction (behaviour, mood, the weather) is possible, long term is impossible(190).
Pseudo-periodicity	Variation in a system may superficially appear to be regular and cyclical, but in fact is not, but nor is it random,	Hormone secretion(194) variations in stride length of normal gait(195)
Power law scaling of events	The distribution of event sizes (e.g. changes in a system or size of an occurrence) is not normally distributed. Rather there are very many small sized events but a few extremely large ones: the distribution follows the inverse power law where the probability of an event is inversely proportional to 1/ the event size raised to a constant power	EEG analysis(196), number of consultations per episode of illness(14), behaviour of hospital waiting lists(197)
Long range correlations	Despite the unpredictable medium term behaviour of the system (see divergence above) long term correlations and patterns are seen in a non-random way	Heart rate variability(198) EEG activation (199)

¹ As an example of the arguments in favour of the evolution of complex systems, this study suggests that a fractal structure combines efficiency in the primary function – transmission of air through the bronchi – with greater ability to maintain that efficiency in the face of change than other structures

Whereas chaos refers to a non-linear system which can be completely defined by a set of difference equations and their starting values, complex systems is a term used to describe non-linear systems which cannot be described so simply, as they consist of many components each with the ability to be influenced by other components of the system. By spreading or inhibiting that influence according to their local conditions and behaviour, elements, or “agents” of the system affect the wider “emergent” behaviour of the whole system. The behaviours of such systems generally share the properties in Table 6 but instead of being modelled mathematically by deterministic equations, they can only be modelled by simulation, using techniques such as agent based modelling(191). Even such models generally isolate the system from its environment and consider it alone, whereas real world complex systems are affected by the context in which they operate.

A simple analogy for complex behaviour is provided by a group of cyclists, whereby each member of the group comprises a complicated combination of cycle and rider. They are broadly similar but have slight differences in technology, attitude and capability. Racing independently from each other in a single sprint the distribution of individual timings would be broadly Gaussian. However once they race as a group, as for instance in the Tour de France, the individual characteristics combine to generate a group effect (the peloton) with some individual outliers (the leaders and stragglers). Within the main group, small changes in one individual (e.g. a subtle swerve) are usually buffered by the group system but may, if the conditions are just right (or wrong!) lead to a wave of instability spreading through the pack, disrupting its shape and trajectory, and possibly causing some of the participants to fall off.

This metaphor is attractive (and computationally tractable, at least on a small scale) for other group situations either at a cellular level (e.g. the interactions within a group of neurons governing a biological system, which take complex inputs and

produce a constantly changing form of stability²) or an organisational one (e.g. the diffusion of a change through an organisation – or more commonly the failure of well intentioned change to take hold.)

Statistical Methods for Complex Systems

The theories of complex systems have resulted in statistical methods, which are computationally intensive and which typically depend on analysis of large datasets, either comprising long time series which are often continuous rather than interrupted or distributions of events within a population. Such methods tend to be derived from first principles and tested first on idealised systems then on actual data. Chaos and complexity have also led to use of new forms of language and meaning which generate potentially useful metaphors – such as the idea of attractors³(192;200) and the much abused “butterfly effect” whereby, depending on the state of all other conditions, a butterfly flapping its wings in China may trigger a hurricane in the Caribbean, although in all probability it won't.

Statistical methods can broadly be defined in three categories: firstly those which seek to identify or characterise the highly specific characteristics of chaotic systems which “are distinguished by sensitive dependence on initial conditions and by having evolution through phase space that appears to be quite random.”(201). These statistics such as fractal dimension and lyapunov exponent typically need large data series (e.g. 10^{4-5} points) for accuracy and could not be attempted on the relatively short data series achievable from a diary study.

² Dynamic stability is a paradoxical term which is useful in referring to the ability of a system to stay constantly changing but always able to react, it contrasts with the concept of equilibrium or perfect balance with stasis.

³ Attractor strictly refers to an area of mathematical space to which a pseudo-periodic non-linear system keeps returning, although never quite entering or exiting the area in the same way; it has been adopted as a way of describing systems behaviour which seems to be “attracted” to certain recurring patterns.

A second set of statistics have been developed to investigate chaos-like behaviour in complex data which aims not to “prove chaos” but to quantify the unpredictability of a system’s behaviour with a numerical value to represent a point somewhere between complete randomness and absolutely reliable repetition or periodicity. These include measures of statistical entropy. One of these, sample entropy was used in the current research and will be discussed in depth later.

Finally a simpler approach is to describe the distribution of one or more characteristics of a system to test for its similarity to a theoretical complex system. Distributions from complex systems are not usually Gaussian, instead they are typically monotonic with the most frequently occurring items also being the smallest, but with a much longer and thicker tail to the distribution than either the Gaussian / normal or exponential / half-life patterns. This means that while there are still very few large items, they are larger than one would expect from conventional probability distributions and occur more frequently. Typically the distribution curve is smoothly concave but when plotted on logarithmic axes becomes a straight line.

Elsewhere, with a colleague, I have demonstrated just such a “power law” distribution in the distribution of numbers of consultations per episode of low back pain in a dataset of over 140,000 episodes of low back pain(14). While similar distributions have been seen in healthcare systems(197) our data are compatible with decision making by individual patients following broadly similar rules regardless of whether they were relatively high and low consulters. Had frequent attenders been a discrete group one would have expected a change in the line of the distribution but from our data, the slope of the distribution was identical between, for example, 1 and 2 consultations per episode and between 30 and 50.

Statistical Entropy

The concept of statistical entropy originated in the mid 20th century with Shannon’s work on information theory. Entropy was broadly defined as the unpredictability of a sequence of items, hence a sequence of repeating symbols – for instance the SOS distress signal, either written or in morse code –has a very low entropy because every “O” will be followed by an “S” and every “S” will either be followed by either an “O” or a pause. In contrast, a random sequence of letters has a high entropy in that

any letter may be followed by any other letter. Clearly a language such as English will have an entropy value somewhere between these two, because while it is infinitely variable, it is not random, with some letter sequences occurring much more frequently than others. Statistically, informational entropy depends on the probability of the next item in a sequence being predictable from those that precede it.

Starting from this point Pincus(202) developed the concept of approximate entropy (ApEn) for a data series based on the probability of two matching subsequences of a time series of length n also matching when the next item is included to make them length $n+1$. Although initially developed for series of around 1,000 points, subsequent work suggested reasonable accuracy with series lengths down to 50 items(194). Richman and Lake proposed a number of improvements to the model and defined sample entropy (SampEn)(203) as “the negative natural logarithm of the conditional probability that two sequences which match for length m also match for length $m+1$ ”. The sample entropy statistic (SampEn) has three parameters: N the length of the series, m the length of the shorter subsequence; and r the permitted tolerance for matches, usually expressed as a fraction of the standard deviation of the normalised data series such that the mean is 0 and the standard deviation is 1.

For both SampEn and ApEn, lower values represent less unpredictability, while higher values represent more unpredictability. There are no calculable normal values and so results must be interpreted relative to either data from control subjects or surrogate data derived from the study data and which possess the same summary statistics (mean, variation, trend etc) but which do not possess the same sequential structure.

Entropy, predictability and health

Common sense, as expressed in lay speech, tends to suggest health as a state of good order while illness is less ordered, indeed the term disorder itself is used in medical discourse. Patients and physicians frequently use metaphors of stability to imply recovery and health.

More recent analysis of physiological systems however suggests that normal behaviour is more random than might be expected from models of smooth orderliness. Indeed studies of several biophysical systems have shown that healthy

systems show less order, more unpredictability and greater statistical entropy than diseased or ageing ones(194;195;204). In fact the complexity of physiological dynamics is not quite so counter-intuitive as the simple notion that more disorder equals better health. Indeed some disease states, for instance atrial fibrillation, are characterised by gross disorder. However studies of longer time series show that in health, a similar pattern of variability to that shown from beat to beat, exists from minute to minute and hour to hour. This would not happen if the variability was truly random (statistically, random variation would average out over larger scales and the apparent unpredictability would fall) but instead represents a form of dynamic stability within which there is constant variation. Indeed in the case of atrial fibrillation this is exactly the case, with high statistical entropy on short term measures, but very low entropy over the long term, whereas healthy individuals show a similar, moderate, entropy score at all scales of time(205).

The term physiological complexity can be used to describe the dynamics of a system in which there is a degree of structure to the temporal variability such that it is neither so simple as to appear ordered to the observer, nor completely random; and in which that structure to the variability exists over several different timescales.

In explaining physiological complexity, Goldberger suggests that the mechanism is multiple interacting feedback systems, in which the degree of interdependence is so high that the system has many different potential responses, and that the objective is to achieve a state which combines adaptability with robustness such that the system can respond to many different challenges and tolerate impairment of one or more parts while remaining functional. He and colleagues talk of a “general complexity-loss theory of disease and ageing”(198).

Complexity Statistics in mental health

Since the early work on chaos theory and biological systems(206)the science of dynamical systems has held out the tantalising prospect of providing understanding of psychological processes and disorders. An early study of patients with bipolar disorder found that the “cycles” in mood were not truly periodic, rather they had mathematical characteristics of the pseudoperiodic changes seen in chaotic systems (207).

Much work has been carried out on EEG recordings in health and disease, typically showing that complexity is higher in wakeful reactive states, and lower in disease (199;208). Recently mania has been identified as a disease state characterised, unusually, by higher complexity in the EEG(209).

Three studies have examined mood changes from the perspective of non-linear statistical methods. Woynshville and colleagues(210) compared daily mood readings in 36 patients and 24 controls using end of day self reported mood readings (average, peak and trough for the day). Patients all reported affective instability (i.e. sudden and marked mood variation) in the presence of either major depressive disorder (12 patients) or bipolar disorder (24 patients). Two thirds also had at least one comorbid dependence disorder in remission and most had an identifiable, if borderline, personality disorder. Three measures of non-linear complexity (mean squared successive difference, fractal dimension and power spectral density) suggested that although the amplitude of the mood swings was higher in patients than controls, the complexity of them was lower.

Yeragani (211) investigated 20 normal controls under 3 treatment conditions: placebo and while taking fluoxetine, an SSRI antidepressant, and pemoline a neurostimulant similar to methylphenidate for 56 days of each. Subjects completed mood self ratings at the end of each day and reported no difference in mood during active treatment. However during treatment with pemoline, complexity of mood, as estimated using the Approximate Entropy statistic, was significantly lower than during either placebo or fluoxetine treatment. The authors suggest this represents a dampening of normal mood variation in healthy controls during treatment with a neuro-stimulant drug.

In a more detailed study of one patient with recurrent major depressive disorder and one control(212), each collecting approximately 1800 data points representing 10, hourly, mood ratings per day for 6 months, Heath and colleagues demonstrated significant differences in several measures of complexity between the case and control. Multiple analytical techniques consistently demonstrated more regularity and less unpredictability in the patient with depression.

There have been no published data using non-linear dynamic methods with medically unexplained symptoms or chronic pain, although Thayer and colleagues have

outlined a possible dynamical framework within which interactions of complex dynamic processes may interact to link together emotion and autonomic responses(213).

Conclusions

The science of non-linear dynamic systems provide a novel perspective on the changing patterns of mood and symptoms. This perspective leads to the hypothesis of illness as a state characterised by loss of complexity and adaptability. It also offers recently established quantitative methods for analysis of datasets which are relatively short and noisy and which allow hypotheses to be tested. This kind of analysis is suitable for accurate series of around 100 points such as those obtained from electronic diaries of symptoms over time.

Chapter 6 Aims and hypotheses

The preceding sections have outlined the relevant literature and theoretical background to the research study. This section describes the study aims and objectives.

Overall aim

The overall aim of this research is to investigate the associations between physical symptoms and emotions or cognitions over time in patients with persistent medically unexplained symptoms.

Specific Questions and hypotheses

Research Question 1 : Are electronic diaries a suitable tool for symptoms research?

The likelihood that electronic diaries are suitable for symptoms research has already been demonstrated in Chapter 3 by analogy to other related research topics.

The empirical section of this thesis reports a study using electronic diaries to track symptoms and psychological variables and data from that were used to test for completeness and accuracy of the data.

Question 2: How do symptoms and emotional states reported by electronic diary vary over time in patients with persistent MUS?

Hypotheses

Symptoms and emotional variables will vary over time, but not consistently so, either within or between individuals.

Symptom ratings will show autocorrelation which can be statistically removed to permit further analysis.

Question 3: What are the concurrent associations between symptoms and mood?

Hypothesis

There will be significant associations between physical and emotional variables.

Associations will differ within and between individuals.

Question 4 *Is there any evidence for consistent sequential relationships between symptoms over time?*

Hypotheses

Time series data from patients with MUS will not show predictable sequential changes to suggest psychosomatic causal sequences.

Time series data will show an increase in concern and / or symptoms prior to GP consultation, in keeping with a previously published model(89).

Question 5: *Do symptom time series data for patients with MUS show signs of loss of complexity?*

Hypothesis

In keeping with Goldberger's model of adaptability as health(195), data from patients with persistently unexplained symptoms will show reduced statistical complexity.

Question 6: *How do patients with MUS describe their condition after viewing their own diary data?*

Hypothesis

Patients with MUS will find the electronic diary, and report of their data, helpful in describing or understanding their experience.

Question 7: *What does it mean to consider MUS as a complex illness?*

Patients with persistent MUS may be described by a complex adaptive longitudinal model of illness rather than as having discrete disorders. The rationale for this and its implications for patient care and research will be elaborated.

Chapter 7 Methods – patients and procedures

Sample

The aim of the sampling process was to obtain a diverse group of patients in terms of age, sex, activity, and medical condition / predominant symptoms. As the study was essentially idiographic and exploratory no statistical assessment of sample size was made. The main study protocol is outlined in Appendix A.

Entry criteria for the study

The target population for the study was adults between the ages of 25 and 65 who were currently regularly experiencing at least three symptoms which could not be adequately explained by organic pathology and affected at least two bodily systems. Permitted symptoms included pain, which could be musculoskeletal, abdominal, pelvic or thoracic, and functional symptoms of the digestive system (dyspepsia, bowel symptoms), cardio-respiratory system (palpitations, throat tightness) or nervous system (headache, dizziness, weakness). The study deliberately used loose criteria for functional disorders as used in epidemiological studies(214) rather than strict definitions such as that for fibromyalgia(215).

Potential subjects were excluded if they had a history of severe physical illness (coronary heart disease, cancer, active inflammatory disease) or depression which was either new or severe. The study required two attendances at health service premises so potential participants were excluded if unable to leave their house without ambulance transport. Patients with depression were excluded if they had recently started an antidepressant for new onset of depressive symptoms, or expressed thoughts of self harm. Subjects taking antidepressant medication for more than three months with no evidence of severe or deteriorating depression were eligible for inclusion, particularly as for many the antidepressant was at least partly prescribed for relief of physical symptoms.

For a symptom to be deemed eligible for the study several features were required. Firstly the participant must have consulted their general practitioner about it and received negative examinations and investigations. Secondly the GP, and any specialist to whom the patient had been referred, must have told them either that the

symptoms belonged to one of the functional physical symptoms or that they were not fully explicable.

Screening for entry criteria / exclusion

Potential participants were screened for entry into the study differently depending on whether they were identified by a physician or self-referred. Patients referred by a physician to the study with a completed referral checklist were contacted by phone to arrange an appointment. A few were referred informally with their verbal consent and they were sent a study information pack by post, which was followed up by a phone call to confirm interest and arrange an appointment.

Potential participants who referred themselves were sent a full information pack and were only contacted again after returning the completed patient checklist and giving written consent to contact their own GP for information. On receipt of this stage of consent the research nurse contacted the individual by phone to screen for entry and exclusion criteria.

Recruitment

Recruitment through physicians

The initial recruitment of patients was by local medical practitioners. Three practices which expressed an informal interest in the study at a very early stage were directly visited to explain the study. A short presentation about the study was also made at a postgraduate meeting. Letters were sent to all practices in the Nithsdale, Annandale and Stewartry areas of Dumfries and Galloway.

The initial information included a brief summary of the study, a copy of the patient information leaflet, and a set of desktop reminders describing suitable patients.

In a later recruitment phase, invitations were also sent to local hospital specialists in gastroenterology, rheumatology, general medicine, gynaecology and rehabilitation medicine.

Where patients were identified by their physician, either in primary or secondary care the doctor was asked to confirm that no contraindications to participation were present.

Recruitment via local media

In response to failure to meet interim recruitment targets after 15 months of the study, a feature about the study was run in the local newspaper and on a local radio station. Interested individuals were invited to contact the researchers directly or discuss it with their general practitioner. The opportunity for potential participants to approach the researcher directly was included in order to avoid extra consultations with GPs at the time of year (February) which tends to be a busy one for practices with winter illnesses. When participants approached the researchers directly, they were asked to consent to a check being made with their general practitioner that there was no reason they should not participate.

Location

The study was carried out in the Dumfries and Galloway region of Scotland. This is the area in which I have practised for 20 years and comprises one medium sized town with a district general hospital and several smaller communities between 30 and 90 minutes travel from Dumfries. The study was not carried out in my own practice area of Upper Nithsdale in order to avoid any conflict of interest.

Participants were seen for the main study visits in local hospital or health centre premises within 30 minutes travel time of their home, and for interim data collection visits were seen in their own homes unless they opted for health service premises.

Ethical Approval

The study was approved by Dumfries and Galloway Local Research Ethics Committee (ref 02/11/07). The initial protocol included only recruitment via GPs, however two subsequent amendments were made, to allow recruitment by hospital specialists and by potential participants approaching the researchers directly.

Study Interviews

First Interview

The first interview for each participant was held in a health service consulting room, either in a hospital, or health centre. The author, as principal researcher, conducted the first part of the interview but left the room, or moved to a corner so as to have minimal influence, while baseline questionnaires were completed. The research nurse

remained in the room throughout the interview. Initial interviews lasted 40-70 minutes. During the interview the principal researcher made brief field notes, both as aides memoire and for further descriptive analysis.

Following introductions and clarification of the nature and purpose of the interview, potential participants were asked to describe and give a brief history of their current medical problems. Where appropriate this was elaborated with further questions in order to give a reasonably full clinical impression of the participant and to ensure that entry criteria were met. The interview style was analogous to a long consultation in clinical practice with the exception that no physical examination was carried out and the results of tests and prescribed medication were taken on the interviewee's word.

Following the clinical history stage of the interview the principal researcher decided whether the patient met the entry criteria. If they did not, they were informed of this and thanked for their attendance. An explanation of why they had not met the criteria was given.

For the remainder, who met entry criteria, the electronic diary was demonstrated and the content of the study information restated prior to completion of the consent procedure. The consent form is included in Appendix B.

When informed consent to participate in the study had been formally obtained, the principal (GP) researcher discussed with the participant which symptoms would be recorded on the electronic diary and asked for the participant's preferred times for the audible prompt for diary entry. He then programmed the diary with to use this information.

While the diary was being programmed the research nurse administered the three baseline questionnaire instruments: Hospital Anxiety and Depression Scale(216), Somatic Symptoms Inventory(217) and the Whitely-7 Index of Somatisation(218). These scales are included in Appendix C to E. The completed measures were then stored by the research nurse until all the data had been collected at the end of the study to maximise blinding during subsequent interviews.

Following completion of the questionnaires, the participant was shown how to use the diary and then completed one, or if wished, two data entry sequences. A printed

diary information leaflet (Appendix G) was given along with a small notebook in which specific events could be recorded.

Data Collection

One week after the initial interview, the research nurse contacted each participant by telephone.

Three or four visits were then made to each participant at approximately 3 weekly intervals by the research nurse. These visits were to the person's own home and lasted between 10 and 30 minutes. The main functions were backup of data from the electronic diary, replacement of batteries, and checking for any problems with the study protocol. No feedback was provided from the diary data and it was not analysed at all, except in some cases to test reliability, during the collection period

Where diary data was lost through technical failure or suspended because of the participant's holiday the study period was lengthened so as to obtain 12 weeks of data. In the case of data loss, the most recent backup of data was restored and the diary recalibrated with maximum and minimum visual analogue scores before being returned to the user.

At the end of the data collection period participants were either visited personally or supplied with a suitable padded envelope to return their diary through the health service's secure internal mail service by handing it in at their GP's surgery.

Towards the end of the data collection period the research nurse left her post and the principal GP researcher made two data collection visits to the last four participants. During these visits any medical discussion was kept to a minimum.

Final Interview

Following completion of the diary, the research nurse organised a final interview with the principal researcher, usually within four weeks of completion of the diary. This was longer for a small number of participants. The final interviews were held in the same places as the first ones.

At the final interview the principal researcher first acknowledged the effort made by the participant in completing the diary then enquired generally about how the participant had fared and whether they had noticed anything in particular. Thereafter

results from the diary were presented as simple line graphs for each variable and as a network diagram of principal interactions.

Because the methods of statistical analysis and presentation were determined to some extent by the nature of the data and the evolving understanding of the researcher, the actual forms of data fed back to participants changed through the study. For example, early feedback used path diagrams with regression coefficients which were then abandoned when it became clear that the nature of the regression equation often left out potentially important links. Later feedback also used some cluster analysis information to group symptoms together to see if this added value to participants.

Following presentation of the data from the diaries the principal researcher and participant discussed the ways this might be understood. For some participants this stage was very brief, for others it provided an opportunity both to state personal beliefs to a professional willing to listen to them and an opportunity to explore past events and interpretations of them. Final interviews were scheduled to run for between 15 minutes and an hour as necessary.

During the final interview the principal researcher took brief field notes similar to those in the first interview, comprising a mixture of short notes, verbatim quotes, and impressions

Transcription and analysis of field notes

Field notes were stored and transcribed at the end of the study into a database. However themes arising from earlier interviews was used in later ones as were evolving explanatory models and ideas such as “all or nothing behaviour”.

During subsequent analysis the original field notes and their transcribed equivalent were repeatedly read and from them, and comparison with published qualitative research, themes were developed and analysed.

It should be stressed that the notes on which this analysis was made were selectively made by the principal researcher and should not be accorded the same weight as taped recordings of interviews which are less susceptible to observer bias.

Chapter 8 Methods – electronic diary

Aims in design

The study diaries were designed to run on standard handheld personal digital assistant (PDA) computers running the palm™ operating system. These were chosen because of availability, portability, maturity of the technology, and the availability of straightforward programming systems for their use. The software was designed for the project with a view to being easy to use, present items sufficiently randomly to discourage unthinking entry and to contain a method of validating entries.

Hardware

Palm PDAs have been in fairly widespread personal & commercial use from the mid 1990's and the technology is relatively mature, combining stability and robustness. Relatively low cost devices were becoming available in 2000 when preparatory work began (approx £120 per item) and the lowest available specification devices were used provided they had a screen area no smaller than 5x5cm.

Nine of the 12 devices used simple replaceable batteries which kept the purchase cost down but carried the risk of battery failure. Users were not expected to change batteries, instead this was done by the researcher at each visit. Three machines added later had rechargeable internal batteries and the users were supplied with a charging cable.

Evaluation of PDAs for illness monitoring suggested that most people are able to use them, however principles of good design(144) were followed such that the diaries could be easily completed using simple screen touches with the stylus. No character writing was necessary.

The devices used were the Handspring Visor (HandSpring, Mountain View, California) and the Sony Clie (Sony Europe, Weybridge, UK). Initially devices were bought from the distributors, after manufacture of the preferred device ceased, additional PDAs were obtained from e-Bay.

Software

Software for data collection was written specifically for the PDAs using the commercial programming tool NSBasic for Palm (NSBasic corporation, Toronto, Canada). Appendix G, the participant information sheet, shows pictures of the PDA screens in use.

At the study registration interview, participants and the researcher chose three key symptoms from a list of 16 which were most applicable to them. These were selected in the diary configuration, along with the person's preferred first name, study ID and, for women, menopausal status.

In daily use, the diary software was designed to be triggered by pressing one of the standard buttons on the device and had a structured programme flow as follows

Stage 1

Welcome screen. If the diary was being used less than 4 hours or more than 48 hours since the last data entry, a screen was triggered warning the user that they had recently used the diary of reminding them to make regular entries.

Stage 2

A sequence of 8 symptom screens, each with a visual analogue scale (VAS) and a next screen button to navigate through the programme. When the stylus was touched close to the VAS line, a mark was drawn on the screen crossing the line at that point. The user had the option of re-drawing this at the time if they wished.

The sequence of screens was randomly determined for each set of entries.

Any attempt to leave a screen without marking the VAS prompted the opportunity to go back and complete the previous step, otherwise once data was entered it could not be changed.

When the user reselected the next screen button, the x co-ordinate of the mark on the VAS was stored as the value. The VAS line on screen was 150 pixels long and so the stored value had a range of 1-150.

Stage 3

One check screen was triggered on approximately 50% of data entries: one of the previous 8 screens was randomly selected by the software to be displayed again.

Stage 4

One current detail screen, with simple tick boxes to indicate if the user had contacted a doctor or nurse, or if a significant personal or health event had happened since last data entry. Premenopausal women indicated whether they were menstruating.

Stage 5

Exit screen, thanking the user for data entry and stating that the device would spontaneously turn off until the next entry reminder.

Before closing down, all data was written to the PDA's memory in a structured format, along with the date and time of the entry and the sequence in which the screens had been displayed

Choice of phrases

The phrasing of the text on the symptom screens was carefully considered before implementation. A standard format was used for physical symptoms as follows "How much have you been bothered by symptom X? Please mark a point on the line between severe symptom X and no symptom X at all."

The VAS itself had end indicators with the words None and Severe, left and right aligned above the ends. Scales were always presented in the same orientation, with increasing severity from left to right. The content of the questions is listed in Table 7 which shows the format for the individual's specific symptoms (here X,Y and Z) and the common variables.

The questions were specifically chosen to avoid the words "fatigue" and "depression". Thus "energy" and "mood" were used and the scores from these inverted for analysis, to represent fatigue and low mood.

Table 7 Questions asked in symptom diary

Symptom	Screen Header	Question	End indicators Left --- Right
X, Y, Z	X	How much have you been bothered by X	None --- Severe
Energy	Energy	How much energy do you have today	No energy at all --- Full of energy
Mood	Feeling up or down	How do you feel generally today?	Fed up --- Really cheerful
Stress	Stress	How stressful are people and things around you?	Not stressful at all --- Very stressful
Concern	Concern about symptoms	How concerned are you about your symptoms	Not concerned at all – Very concerned
Anxiety	Feeling Calm or on Edge	How anxious do you feel	Really calm – Totally on edge

Entry prompts and data storage

Reminders to enter data were set for two times each day using the PDA's inbuilt calendar application. When the time occurred the device sounded three beeps, and repeated this after 5 and 10 minutes. Thereafter no further alarm sounded until the next time. Times for data entry were chosen by each participant to fit their daily routine.

Data were collected off the devices by connection to a laptop computer or, in the case of the Sony devices, by backing up the data files onto a memory stick using the software pre-programmed into the device. Backed up data were then stored on another computer for subsequent processing and analysis.

Piloting the Electronic Diary

Prior to the study two healthy volunteers used the diary for 1-2 weeks to test reliability and acceptability before two patients of my practice with MUS agreed to pilot data collection for a dull 12 weeks. Both subjects completed more than 80% of recordings and showed good test-retest reliability from comparison of check screen scores with original values. Following their comments, the wording for the mood question was changed from depression to "feeling up or down" as described above.

Review of the diary during the study

After the first six participants had completed their data entry, the electronic diary method was reviewed. No changes were suggested by participants who were happy

with the method, and preliminary calculations of test-retest reliability (as described below) were made to ensure adequate data quality.

Chapter 9 Methods – principal methods of data handling and analysis

Introduction

This section describes the methods of data processing and analysis after the handheld computer data had been transferred to a standard personal computer. Individual methods are not described in detail here, rather they are considered along with the relevant results. Instead, this section contains a broad outline of the analyses used to give an overview of the subsequent detailed analysis.

Data Preparation

Once data collection for an individual was complete the data was saved to a computer file in the standard palm™ database format (.pdb). Data was prepared for analysis in the following steps:

Stage 1: conversion of data file

Data files were converted from palm PDB files to a standard computer format (tab delimited text) by the computer programme PDB Converter (www.mverive.com/pdbconverter.htm) and imported into a spreadsheet (Excel 2000, Microsoft corporation)

Stage 2:Preparation for test-retest reliability testing.

For entries during which validation screen had been presented (approximately 50% of entries), the matching original value was copied alongside the check value to create a pair of original and retest values.

Stage 3: Preparation for analysis

Ratings for mood and energy were reversed to obtain scores for depression and fatigue respectively. Initial calibration results were removed (before any PDA was used, the software was calibrated with a series of minimum and maximum values). Where re-calibration entries were made if a PDA had been replaced in mid study these were also removed.

Stage 4: Saving and storing data

The processed data was saved as an excel file to be used by SPSS and other analysis software. One file was saved for each subject. Data was kept on one computer and

regularly backed up to other secure media. No data stored electronically contained patient identifiable material.

Statistical software used for analysis

Several different computer software programmes were used in the analysis.

Simple time series plots of data were generated in Microsoft Excel. Most of the final conventional statistical analysis was carried out in SPSS 11.5 for windows (SPSS inc, Chicago). Some analysis required unconventional procedures run repeatedly for multiple variable combinations on separate data files and for these the R statistical system(219) and hand-coded scripts in the Python™ language(220) were used. The choice between R and python for individual procedures was made on an ad hoc basis, Python is the more verbose and accessible language of the two, with greater flexibility for wider data handling procedures than just statistical processing, it was used for situations where R did not offer a module to carry out a specific task. R and Python scripts are included in Appendix H -J.

Intra subject measures

Plotting charts

For each subject the data were first plotted as a set of eight single variable line graphs. No distinction was made between time of day in these charts.

This technique produced a way of easily visualising data and assessing time trends, unusual episodes, and an approximate idea of the variance in any measure.

Data quality

Data were visually inspected for unexpected values: particularly missed individual data entries with no value and recalibration sets made up of the maximum and minimum values.

Test-retest reliability was estimated using the intraclass correlation coefficient on pairs of check values and their matching data item.

Missing data where no entry had been made at a scheduled time were deemed lost and no attempt was made to impute these to make complete series. Instead measures of data completeness were developed and compiled during analysis. Data was

filtered before final analysis to only include data from days with two diary entries. The rationale for this is described in more detail in Chapter 12.

Autocorrelation

Analysis of autocorrelation was carried out in using a script for batch operation in the R statistical system. As differenced data were subsequently used in analysis, autocorrelation analysis was also carried out on differenced data. Cross correlation between variables was not carried out as it is highly sensitive to the presence of autocorrelation and thus difficult to interpret (183) p159).

For subsequent analysis, autocorrelation was removed by differencing the data (replacing the original sequence of length N with a sequence of length N-1 representing the difference between each original data point and the one preceding it). Details of the methods for testing, reporting and dealing with autocorrelations are in the section on preliminary time series analysis.

Correlation matrix

The correlation between symptoms at each time point was calculated in SPSS to yield a correlation matrix for each individual. Three possible sources of statistical bias were considered which could affect these correlations. The first of these was the effect of autocorrelation, the second was time of day, as subjects recorded data twice most days and correlation of fatigue and pain may have been more to do with the time of day, than a direct link. The third factor is long term trend, if two symptoms both increased or decreased through the time of the study for independent reasons, they would still appear correlated.

To address these problems, a number of additional correlation matrices were generated and considered, for instance using differenced data or partial correlations adjusting for time of day and date, and the combination of both of these. The rationale for choice of matrix is explained fully in Chapter 11.

Changes in relation to consultation

The diary was designed to capture when participants had consulted a doctor and could thus be used to test the hypothesis that symptoms and concern increased before a consultation and that they were subsequently reduced after it.

In both cases, average values for variables were compared: this provided a crude measure of the change in the mean value at the time of the consultation. Three arbitrary intervals were used: The pre-consultation period was the day of the consultation and the two preceding ones; the post consultation period was the three days after the consultation; and the baseline period was the week from 4-10 days before the consultation.

The diary was not able to differentiate between elective consultations and ones triggered by worsening symptoms.

Timing of interactions and Granger Causality

Granger causality is estimated by estimating the regression coefficient r of an autoregressive model of one variable Y then recalculating this with lagged values of the other variable X added to the equation. If an F-test between the two coefficients shows a significant difference, then X is said to granger-cause Y . In practice such relationships are frequently reciprocal.

Analysis of granger causality was carried out using the `grangert.test` procedure of the `MSBVAR` module in the R statistical package(221). The code for this script is in Appendix I.

Pooling between subjects

Meta-analysis

Data from individual subjects was compared by meta-analysis of correlation coefficients. This method was chosen in order to harness the statistical power of the large within-subject sample size relative to the small between-subject sample size. Because subjects recorded different symptoms from each other, correlation coefficients between each symptom pair were used, rather than more complicated regression equations for which different individuals would have different dependent variables.

Meta-analyses of correlation coefficients was carried out using the variable effects method of Hunter and Schmidt(222) as delineated by Field(188). This takes account of sources of error, in this case indices of data accuracy, in both the weight given to different individuals in the meta-analysis, and also the confidence intervals applied to them.

Results for a sample of analyses were validated by checking results against published meta-analysis software using a similar method(223): meta-analytic correlation coefficients varied between methods by <0.02 . Because of the judgement required in estimating the effect of artefacts in the data, it was possible to obtain a wider range of results according to the weighting given to sample size and accuracy. The method chosen adhered strictly to the process detailed by Hunter & Schmidt and a weighting was used which gave conservative values for the meta-analytic correlation coefficients. The software was hand-coded to ensure that a batched series of meta-analyses could be run from data files already stored on disk (approximately 80 separate analyses), without re-entering data and risking its corruption. The computer scripts for preparing and grouping the correlation coefficients, carrying out the analysis and making forest plots are shown in Appendix H. The detailed steps of the meta-analysis are described in Chapter 13 along with the results.

Each meta-analysis was used to derive a pooled effect size with 95% confidence intervals and with credibility intervals used as a measure of heterogeneity(222;224). All results were presented in tabular format and as sorted forest plots.

Non-linear Analysis

For non-linear analysis, sample entropy was calculated at the level of the individual person-variable series. For comparison, surrogates were derived for each series using the Surrogates(225) programme in the Tisean(226) suite of non-linear times series analysis software. The technique is described in more detail, along with the results, in Chapter 16

Techniques considered but not used

Multivariate Linear Regression

Multivariate linear regression, to assess associations between variables, was examined during the study using both differenced data, and including previous values as additional independent variables. Regression between variable pairs was used, both as the correlation between variable pairs (which is equivalent to univariate regression of one of the pair on the other) and in the analysis of granger causality.

While multivariate regression was capable of producing statistically significant models its results are not described for four reasons.

Firstly regression techniques generate a single solution which is mathematically the best fit, but which can be too parsimonious to describe the complex relationships between variables. For instance two closely correlated variables (such as stress and anxiety) may both be correlated with a physical symptom, but the optimal regression solution may only include one. This does not rule out an effect of the other, although it is excluded from the regression model.

Secondly the solution selected by the regression model is only one of many possible models. The simplification of a complex web of interacting symptoms into a sparse regression model probably misrepresents the complex nature of the problem of MUS.

Thirdly, regression methods depend on assumptions of independence of variables and of equality of variance which are certainly missing from raw data. While stepwise models adjusting for preceding values of a variable and additional factors such as time or long term trend are possible and have been used elsewhere(49), the other reasons listed here mitigated against using this approach.

Finally, regression was discarded as a method of analysis for this study because of the difficulty in making comparison between individuals with different combinations of symptoms. As few subjects shared the exact combination of chosen physical symptoms with another, regression models would be invalid for comparison between subjects.

Time Series Models e.g. ARMA

Conventional time series analysis such as the ARMA family of models and more sophisticated multivariate models such as multivariate vector autoregression are not reported for the following reasons:

Formal interrupted time series analysis requires data to be stationary and with sampling at regular intervals. In this study, times for data sampling were chosen by participants for their convenience, but some leeway was allowed. This breaks the regularity to some extent, although could perhaps have been allowed by treating the two time entries as ordinal values such as “early” and “late”.

Furthermore techniques such as ARMA seek an optimal solution for a single model. As with multivariate regression, this means that solutions tend to be both parsimonious and unique to the individual and hence less amenable to comparison

between individuals. While there have been reports of pooling of ARMA models by meta-analysis(227) this is not a common technique.

Chapter 10 Results - Description of Participants

This section describes the participants and the processes by which they were recruited to the study. It includes the symptoms chosen for the diaries and the results of baseline questionnaires.

Recruitment process

Fifty four individuals were identified as possible participants in. Sixteen responded to invitations from their doctors (13 from GPs, three from hospital specialists) and 38 approached the researchers directly after hearing about the study in local media.

Of these, 27 (50%) were eligible and consented to take part in the study. One of the 27 elected to withdraw from the study after approximately 5 weeks without specific reason: data from this participant were not used in any analysis. This left 26 to complete the data collection. Figure 2 illustrates the flow of participants through the study.

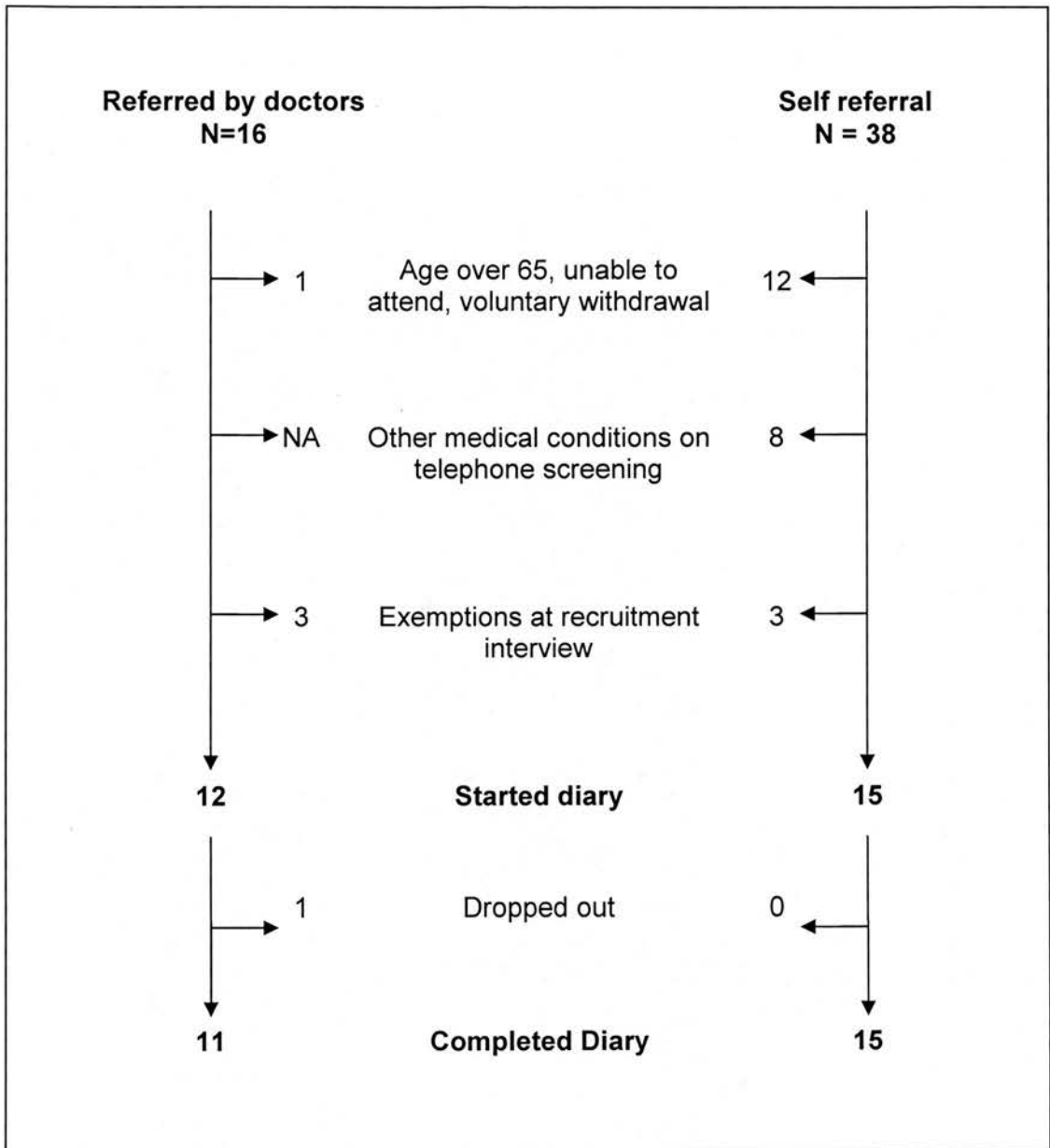
Exemptions

Eight people who had referred themselves for the study and expressed a willingness to discuss their condition by telephone with the research nurse were ineligible, either due to recent new treatment for depression (2 cases) or concurrent physical illness (6 cases: ulcerative colitis(2),stroke, renal colic, myopathy and polymyalgia).

Six potential participants were excluded at the time of the recruitment interview.

Three were thought during the recruitment interview to have “explained causes” for at least some of their symptoms (migraine, cervical spondylosis, inflammatory bowel disease), one had severe depression, one had a pain disorder confined to one limb and one was currently awaiting further medical investigations.

Figure 2 Flow of participants through study



Participants

The 26 participants who completed the diaries ranged in age from 29 to 59 (mean 45, SD 8). Twenty two (84.6%) were female. Eleven (42.3%) were currently in, or actively seeking, work either as a paid employee or as a volunteer with a regular commitment. Either in the past or currently, four had held professional jobs (teaching, nursing) and a further nine had either been educated to degree level or held skilled non-manual posts. Those not working were divided into two arbitrary groups, according to the level of their self-reported dependency on others for simple day to

day tasks: 8 were deemed relatively independent and 7 relatively dependent.

Eighteen were married or in a settled relationship, four were separated or divorced and four were currently single.

Physical Symptoms

Table 8 shows the breakdown of participants by gender, age group and also by principal symptom and working status. For this table very broad symptom categories were used, with “Pain” used for musculoskeletal pain including fibromyalgia and “GI” referring to a range of gastrointestinal complaints including painful irritable bowel syndrome. As participants were required to have several functional symptoms these broad categories are very generalised.

Table 8 Breakdown of participants by age, gender, principal symptom and occupational status

Gender	Age	Main Symptom	Working Status	Participant
Female	21-44	Fatigue	Y	42, 55
			N	49, 53
		Pain	Y	52, 65
			N	56, 63
		GI	Y	46, 54, 60
			N	-
		Other	Y	-
			N	44
				-
	45-65	Fatigue	Y	-
			N	57, 61
		Pain	Y	58
			N	32, 35, 64
		GI	Y	-
N			45, 51	
Other		Y	36	
		N	38	
			-	
Male	21-44	Pain	Y	48
				-
	45-64	Fatigue	Y	37
			N	33
		Pain	N	40

Table 9 lists the specific symptoms chosen by participants to record at each diary entry. The commonest symptoms (with number of participant recording this) were muscle pain (15), headache (12), joint pain (12), bowel trouble (10) and abdominal pain (9). Less commonly chosen symptoms included indigestion (5), nausea (3) and throat tightness (3). Chest pain, pelvic pain and weakness were each selected by two participants and back pain, numbness and palpitations by one.

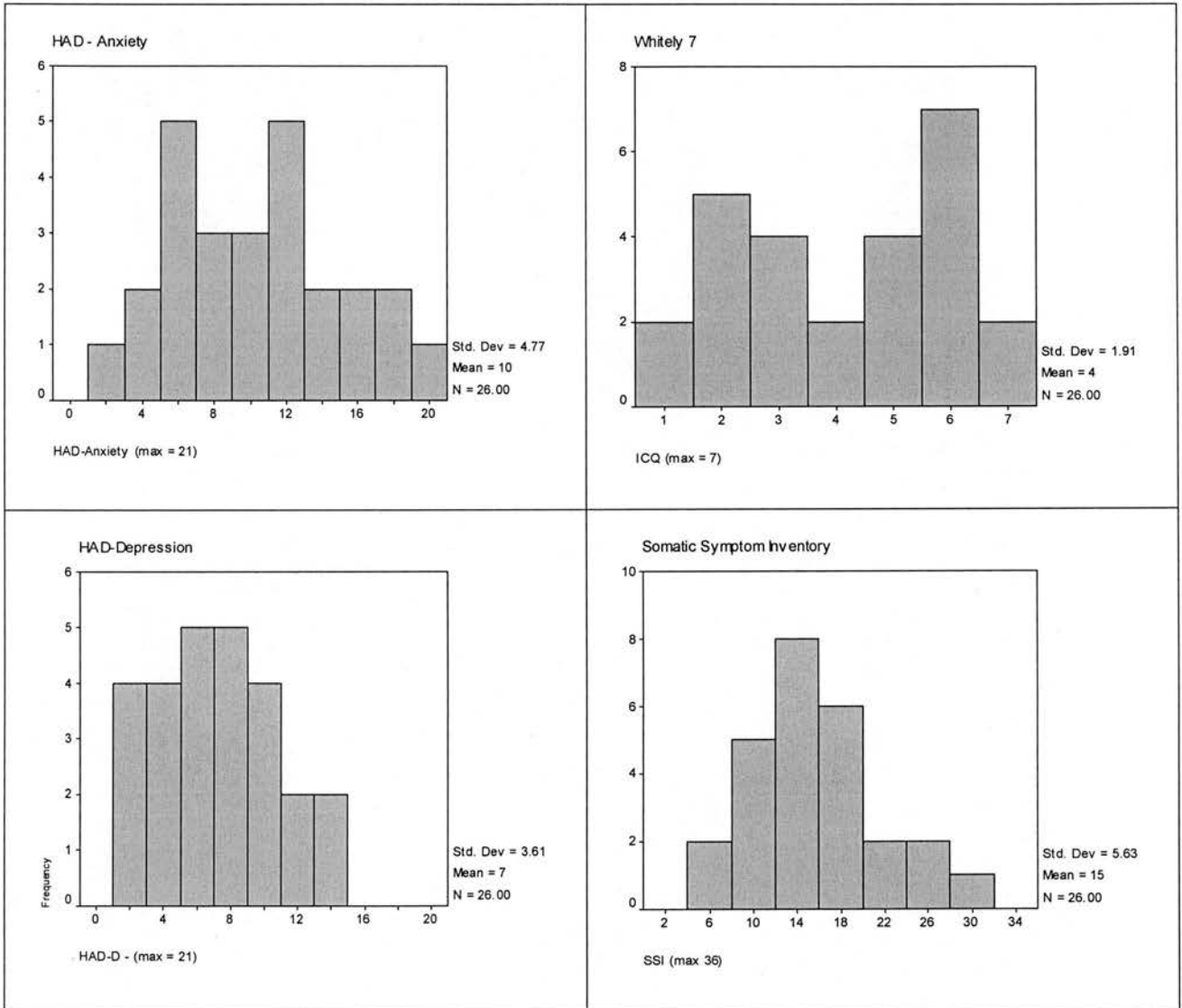
Table 9 Symptoms recorded by each participant

ID	Age	Symptoms		
32	53	Joint Pain	Headache	Bowel Trouble
33	53	Muscle Pain	Bowel Trouble	Nausea
35	56	Muscle Pain	Bowel Trouble	Headache
36	58	Headache	Joint Pain	Bowel Trouble
37	57	Headache	Muscle Pain	Abdominal Pain
38	53	Muscle Pain	Pelvic Pain	Indigestion
40	58	Muscle Pain	Tight Throat	Indigestion
42	43	Headache	Indigestion	Muscle Pain
44	41	Joint Pain	Headache	Bowel Trouble
45	47	Abdominal Pain	Indigestion	Muscle Pain
46	35	Abdominal Pain	Joint Pain	Chest Pain
48	39	Joint Pain	Muscle Pain	Abdominal Pain
49	34	Muscle Pain	Joint Pain	Bowel Trouble
51	46	Abdominal Pain	Bowel Trouble	Headache
52	41	Pelvic Pain	Joint Pain	Nausea
53	29	Muscle Pain	Numbness	Nausea
54	36	Muscle Pain	Tight Throat	Abdominal Pain
55	42	Joint Pain	Palpitations	Indigestion
56	39	Joint Pain	Weakness	Headache
57	49	Muscle Pain	Weakness	Bowel Trouble
58	50	Back Pain	Joint Pain	Bowel Trouble
60	42	Headache	Tight Throat	Abdominal Pain
61	46	Headache	Bowel Trouble	Joint Pain
63	39	Muscle Pain	Headache	Abdominal Pain
64	59	Muscle Pain	Chest Pain	Abdominal Pain
65	44	Joint Pain	Headache	Muscle Pain

Baseline Measures

Figure 3 shows the distributions of baseline scores on the four questionnaires measuring anxiety (HAD-A), depression (HAD-D), illness concern (Whitely 7) and number of recalled physical symptoms (Somatic Symptoms Inventory).

Figure 3 Histograms of baseline questionnaire scores



Correlation coefficients were calculated between questionnaire scores using non-parametric methods as the Whitely 7 scores were not normally distributed (Shapiro-Wilk statistic = 0.902, $p = .02$). The two HAD subscales were significantly correlated with each other (Spearman's rho = .604, $p < .001$) and with the Whitely 7 (rho = .612,

$p=.001$ and $\rho = .431$, $p=.026$ for HAD-D and HAD-A respectively). Somatic Symptoms Inventory was not significantly correlated with the other questionnaires.

Analysis of variance of questionnaire scores with activity status (working, independent or dependent) showed significantly higher numbers of reported symptoms in the dependent group (20 items) compared to either the working (13) or the independent (14) groups ($F_{(23,2)}=4.65$, $p=.02$).

Conclusion

The study participants were a heterogeneous group of patients with a variety of principal symptoms and a range of scores on simple scales for mood disorders and illness cognitions whose lives range from active participation in work to a high level of disability and dependency. Approximately half of the group scored above usual screening threshold scores for the baseline tests.

Chapter 11 Results - Quality of electronic diary data

This section addresses research question 1: the suitability of electronic diary data for symptoms research. In doing so it describes the performance of the electronic diary in terms of how much it was used, how well participants kept to the schedule both within days, and across days, and demonstrates the reliability of data entry as judged by the built in facility for calculating test-retest reliability. It concludes by deriving a single index of data validity for each participant and examining the problem of reactivity, whereby the act of completing the diary itself might have affected the results.

Performance of the electronic diary

At recruitment, none of the participants had experience of using PDAs but all were able to demonstrate use of the electronic diary after brief explanation and assistance. No potential participant withdrew after being shown the diary.

One subject commenced the study but withdrew after 5 weeks. The reason given was personal rather than related to the study methods. All other participants completed the study for 12 -13 weeks.

In total 4008 sets of data were collected in entries on 2314 days during the study. The mean number of entries per individual was 154, range 69-154. Only two participants had less than 120 entries (69 and 111 respectively, the former having lost data due to technical failure near the end of the study when it was not possible to allow extra time for data collection).

Adherence to data entry schedule

Individuals chose times to be reminded to enter data which fitted with their daily routine. The most common pattern was to have the first alarm at a point between 12.00 and 14.00 and the second between 18.00 and 20.00. Several participants chose a later entry time, at the end of the day between 22.00 and midnight. Entries before mid-morning were discouraged at the initial interview in order to allow enough of the day to elapse for differences to appear between days, although one participant chose 09.00 as most appropriate for themselves. In addition, one participant (subject 37) chose to add a morning reading to his scheduled later data entries on most days,

although he was not instructed to do this. Where entries were made in the two hours after midnight these were recoded as occurring at 23.59 on the previous day.

It was clear that some participants adhered more closely than others to the data entry schedule and adherence was measured qualitatively and quantitatively.

Qualitative patterns of data entry adherence

For each participant, a histogram plot was made of the time of each electronic diary entry. Data were grouped in bins of two hours duration. Visual inspection of the plots was used to derive a classification into three groups:

group 1, excellent adherence:

In this group, the histogram showed two distinct narrow columns, which between them contained over 90% of all entries

group 2, good adherence:

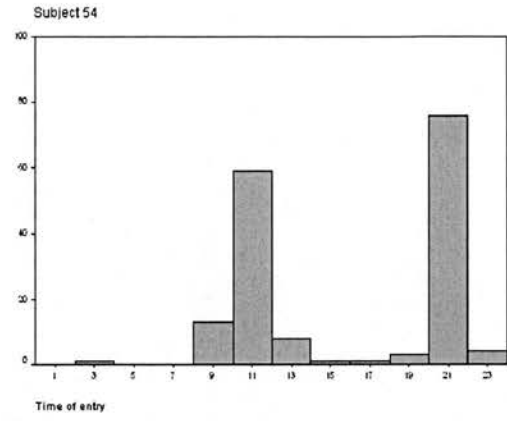
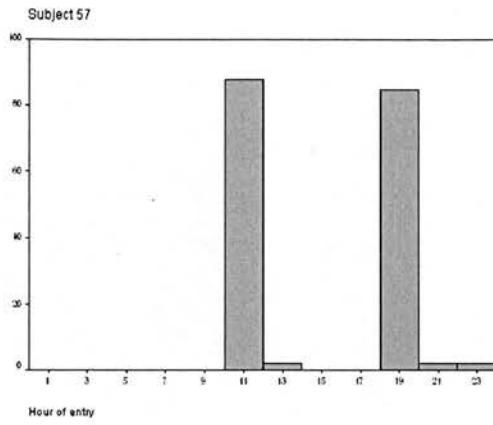
Histograms in this group showed two distinct columns, which were either broader than those in group 1, or which together contained less than 90% of all entries.

group 3, poor adherence:

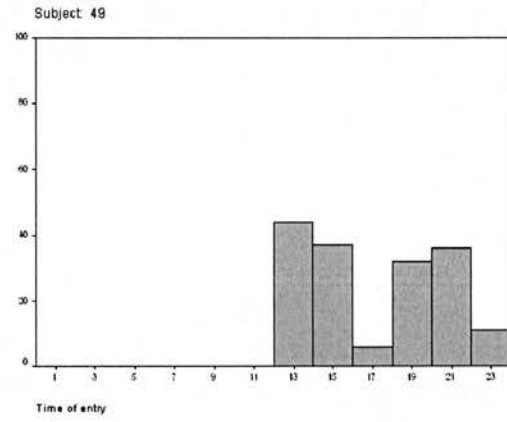
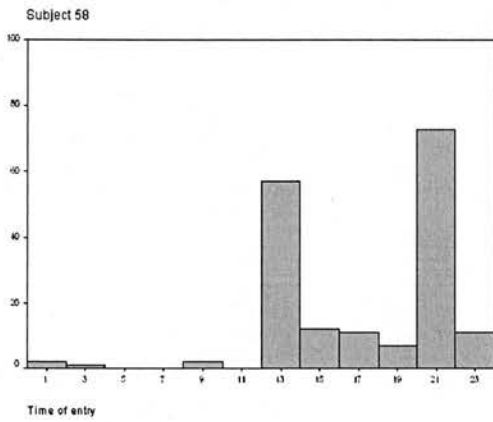
In this group the columns were less distinct and had more entries between the columns than those in groups 1 or 2.

Figure 4 shows typical examples of each pattern. Although relatively poor in adhering to the timing of data entry schedule, even participants in group 3 showed a recognisable pattern of data entry. In view of the overall adequacy of timing of data entry, no participants or data were excluded from analysis.

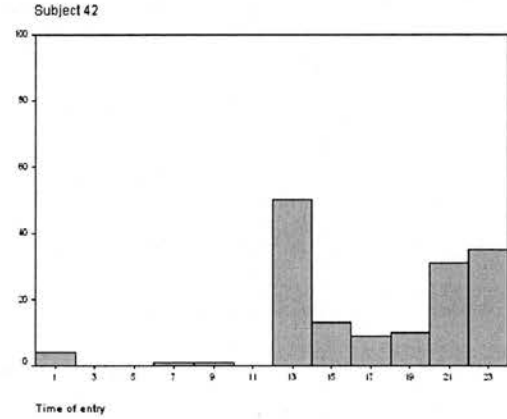
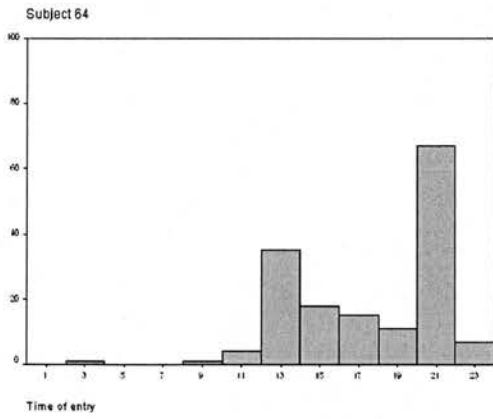
Figure 4 Examples of histograms of entry time



Group 1: excellent adherence to timing regime, narrow peaks with little data outside peaks



Group 2: good adherence to timing regime -broader peaks or more entries outside them



Group 3: poor adherence to timing regime; less distinct peaks

Quantitative measures of data entry adherence

In addition to the qualitative assessment of the pattern of diary entries, a number of quantitative measures of adherence to the diary schedule were estimated in order to assess the completeness and validity of the data.

Calculation of crude data entry rate

For each participant the number of diary entries was counted and the difference between the start and finish dates for each individual was used to estimate the maximum possible number of entries. From this a crude data entry rate was calculated as the number of entries divided by twice the number of days (to account for the two entries each day) including the start and finish dates.

Adjusting for missing data

In some cases data was missing either through technical failure of the diary (4 subjects, 113 days) or due to suspension of data collection while on holiday (2 subjects, 21 days). For these periods of seven days or more during which no data were entered, the maximum number of possible entries was reduced accordingly. Three participants (42, 48 & 64) had brief hospital admissions during the study, with tachycardia, wisdom teeth extraction and abdominal pain respectively, but missed data in these periods, each of less than a week, were not excluded in the same way as holidays or data loss.

After excluding data loss and holidays, a simple data entry rate was calculated for each participant as actual entries divided by twice the number of possible entry days. Values for this ranged between 73% and 107%, with 18 of 26 participants (71%) scoring 90% or more. In two cases the value was greater than 100% owing to an extra, third, entry on some days.

Adjusting for excessive data entry on some days

To counter the effect of extra data entry on some days, an adjusted data entry rate was calculated as follows

$$\text{Adjusted data entry rate} = \frac{(\text{Actual entries} - \text{extra entries})}{\text{possible entries}}$$

The mean adjusted data entry rate was 88% with a range of values between 66% and 100%. Fourteen participants (54%) achieved rates of 90% or more; six (23%) achieved 80-89%; and three (12%) 70-79% and 66-69% respectively.

Adjusting for incomplete entry on days

Because much of the subsequent analysis depended on the temporal characteristics of the data such as diurnal variation, the number of days in which two (or more) entries were made was calculated. The paired data entry rate was calculated for each subject as follows

$$\text{Paired data entry rate} = \frac{\text{days with two or more entries}}{\text{possible entry days}}$$

The mean paired data entry rate was 80.6% (range 41% to 100%). Twelve (46%) participants achieved rates of over 90%; four (15%) 80-89%; three (12%) 70-79%; four (16%) 60-69% and three (12%) 50-59%. In total there were 1758 subject-days with a pair of diary entries, from a possible 2180. The number of paired data days per participant ranged from 27-89 with a mean of 68 and median of 75.

This subset of data, representing only those days on which two entries were made, will be referred to subsequently as daily paired data and used in the time series analysis. Detailed results for each individual are shown in Table 10

Comparison of qualitative and quantitative assessments of data entry.

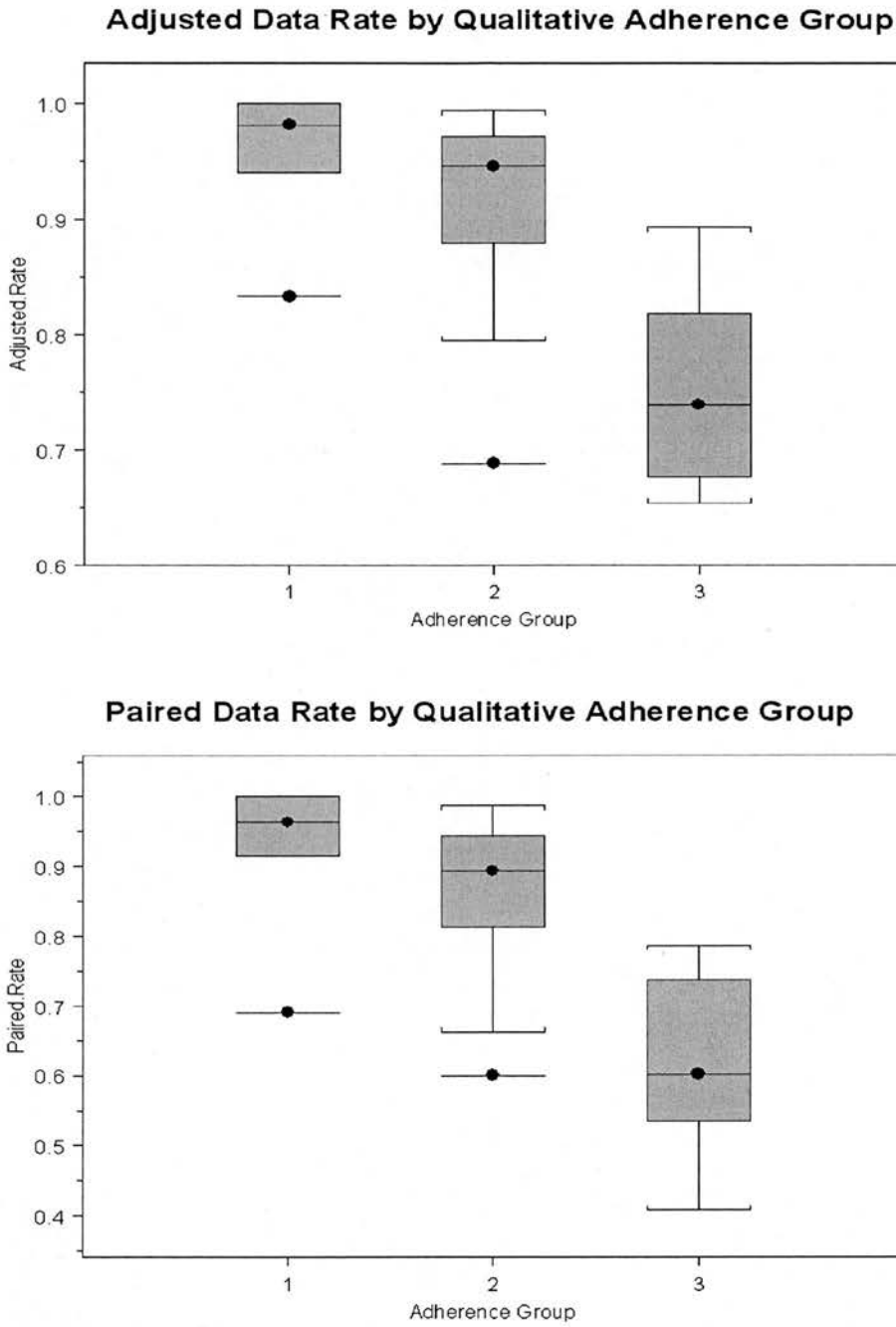
To compare the qualitative and quantitative measures of adherence to the diary entry schedule, a univariate analysis of variance (ANOVA) was performed between the qualitative adherence pattern group and data entry rates. This is shown in Figure 5 Comparison of Data Entry Rates by Qualitative Schedule Adherence and was statistically significant for both Adjusted Data Rate ($F_{2,23} = 13.2, p < .001$) and Paired Data Rate ($F_{2,23} = 14.05, p < .001$). In both cases pairwise comparisons showed no significant difference between groups 1 and 2 but a significant difference between them both and group 3. The two quantitative measures, Adjusted Data Rate and Paired Data Rate were strongly correlated ($R\text{-Squared} = 0.97, p < .001$).

Table 10 Data entry rates for each participant

ID	exempt days	Max days	Max possible	Actual	Simple Rate	missed days	full days	half days	triples	Paired Rate	Adjusted Rate
32	7	76	152	111	0.73	7	31	38	11	0.41	0.66
33	0	83	166	158	0.95	-	74	9	1	0.89	0.95
35	0	83	166	136	0.82	6	55	22	4	0.66	0.80
36	0	85	170	167	0.98	2	76	7	8	0.89	0.94
37	0	86	172	184	1.07	9	64	13	43	0.74	0.82
38	0	84	168	144	0.86	2	58	24	4	0.69	0.83
40	0	83	166	160	0.96	3	76	4	4	0.92	0.94
42	0	84	168	154	0.92	-	66	18	4	0.79	0.89
44	0	84	168	169	1.01	-	81	3	4	0.96	0.98
45	0	82	164	164	1.00	-	79	3	3	0.96	0.98
46	20	63	126	122	0.97	6	46	11	19	0.73	0.82
48	0	91	182	163	0.90	5	74	12	3	0.81	0.88
49	0	83	166	166	1.00	-	82	1	1	0.99	0.99
51	14	94	188	146	0.78	21	50	23	23	0.53	0.65
52	0	86	172	164	0.95	3	76	7	5	0.88	0.92
53	0	84	168	170	1.01	-	84	-	2	1.00	1.00
54	0	83	166	166	1.00	-	81	2	2	0.98	0.99
55	38	92	184	139	0.76	18	54	20	11	0.59	0.70
56	0	85	170	165	0.97	-	79	6	1	0.93	0.96
57	0	89	178	179	1.01	-	89	-	1	1.00	1.00
58	0	89	178	176	0.99	-	84	5	3	0.94	0.97
60	0	84	168	165	0.98	-	79	5	2	0.94	0.97
61	18	45	90	69	0.77	10	27	8	7	0.60	0.69
63	37	88	176	171	0.97	2	81	5	4	0.92	0.95
64	0	99	198	159	0.80	11	61	27	10	0.62	0.75
65	0	95	190	141	0.74	8	51	36	3	0.54	0.73

Max Days number of days on which data entry possible (allowing for exempted days)
Max Possible maximum expected entries – allowing 2 per day including start & finish days
Simple Rate Actual / Max Possible (see text) ; Missed Days: Non exempt days on which no entry made
Full Days: Days with two or more entries made; Half Days: Days on which only one entry made
Triples Days on which three entries were made. NB these are a subset of Full Days.
Paired Rate Proportion of days on which two or more entries were made
Adjusted Rate Proportion of possible entries actually completed, ignoring third data entry on any day

Figure 5 Comparison of Data Entry Rates by Qualitative Schedule Adherence (group 1 had best adherence to the data entry schedule)



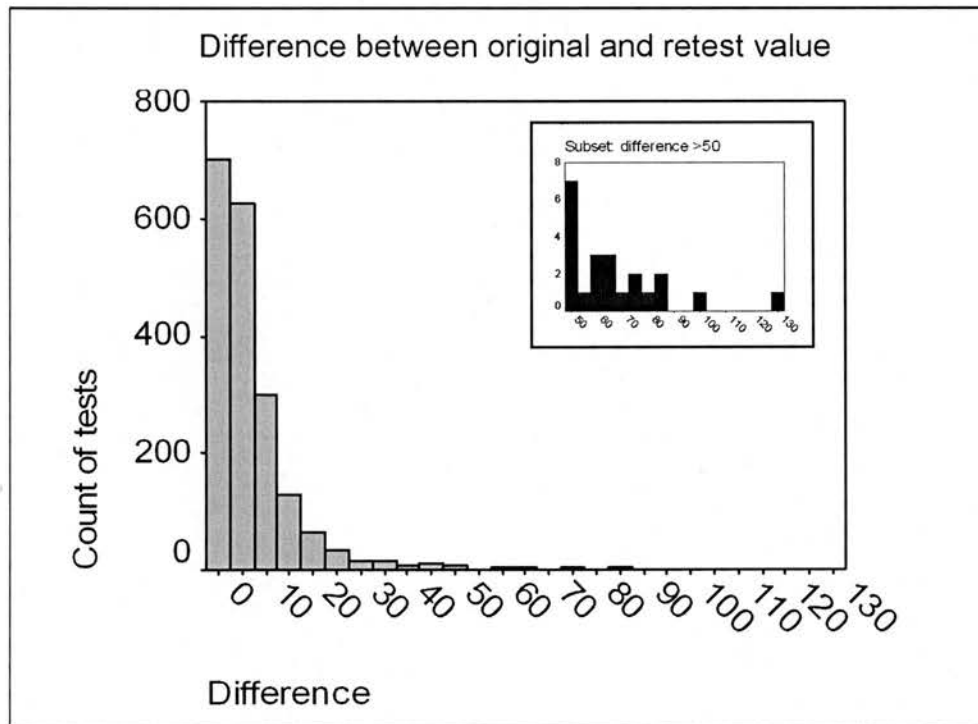
Reliability of data entry

The reliability of data entered in the diaries was checked by comparing check values with the matching original entry. The diary software algorithm was designed to display a check screen in approximately 50% of cases. Check questions were

available for 1928 (48%) of entries and were randomly and approximately equally distributed between variables.

The mean absolute difference between original and check values was 7.2 (SD10.02). As the scale recorded the pixel value from the electronic diary screen and was calibrated from 1-150 this represented 4.7% of the total scale range. The difference was less than 15 (10% of the possible data range) in 1705 (88.3%) readings and less than 8 (5% of possible range) in 1244 (64.4%). There were 22 data entries in which the difference was over 50. Figure 6 shows the distribution of values, with an inset histogram of the values over 50.

Figure 6 Histogram of the absolute difference between check and matching original entries



Test -retest reliability

Formal comparison of the between test variability was carried out by calculating the intraclass correlation coefficient (ICC). This was used to test for differences in reliability either between different items, or between individual participants.

Comparison between variables

Comparison of reliability data between variables recorded, with pooling of individuals, was carried out for the five variables common to all participants. Individuals' chosen physical symptoms were all grouped together. The results are shown in Table 11 which lists the Average of Raters Intraclass Correlation Coefficient, which is equivalent to Cronbach's alpha. All variables have similar ICCs except for fatigue which is significantly lower but still at a level (0.93, 95% confidence intervals 0.92-0.95) which represents excellent reliability.

Table 11 Intraclass correlation coefficient for variables

	N	ICC	(95% CI)
Symptoms	698	0.9795	(0.976 - 0.982)
Fatigue	241	0.9343	(0.915 - 0.949)
Stress	242	0.9819	(0.977 - 0.986)
Depression	253	0.9783	(0.972 - 0.983)
Anxiety	241	0.9743	(0.967 - 0.980)
Concern	253	0.9736	(0.966 - 0.979)

Comparison between individuals

Comparison of reliability data between individuals, with pooling of variables, showed greater differences although these were still modest and overall test-retest reliability was very good with a mean Single Rater Intraclass Correlation Coefficient of 0.938 (range 0.687 – 0.988) and median of 0.963. Table 12 shows the ICC for each participant. There appears to be a degree of heterogeneity between individuals as the 95% confidence intervals for individuals included the group median for only 9 participants, lying below it for 6 and above for 11.

Table 12 Intraclass Correlation Coefficient between individuals

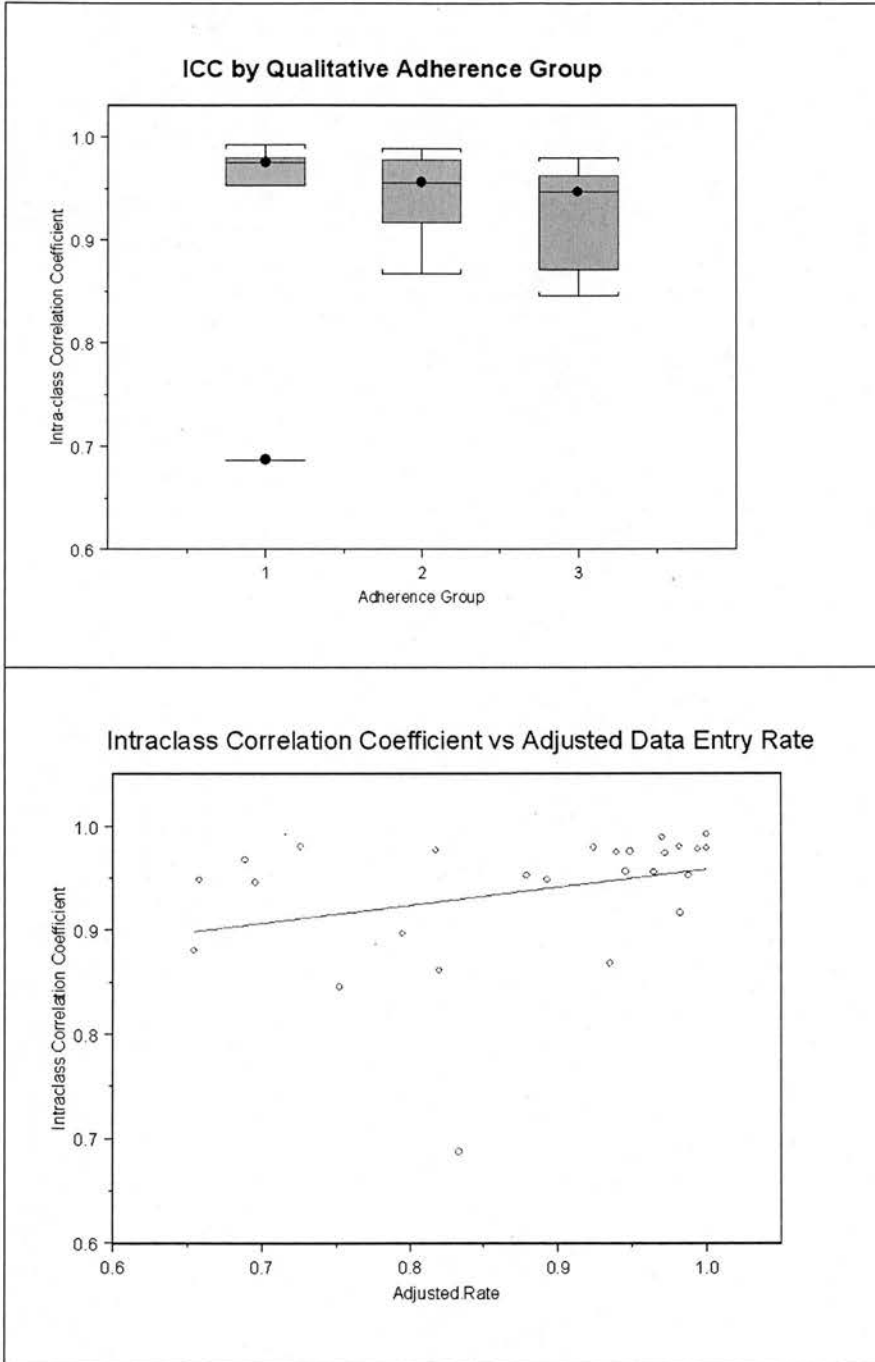
ID	N	ICC	(95% CI)
32	54	0.948	(0.912 - 0.970)
33	69	0.9558	(0.930 - 0.972)
35	64	0.8675	(0.791 - 0.917)
36	83	0.8965	(0.844 - 0.932)
37	86	0.8611	(0.795 - 0.907)
38	71	0.6871	(0.542 - 0.793)
40	81	0.9749	(0.961 - 0.984)
42	68	0.9481	(0.917 - 0.968)
44	77	0.9164	(0.872 - 0.946)
45	76	0.9798	(0.968 - 0.987)
46	64	0.9765	(0.962 - 0.986)
48	76	0.9798	(0.968 - 0.987)
49	89	0.9777	(0.966 - 0.985)
51	71	0.8801	(0.924 - 0.814)
52	76	0.9792	(0.967 - 0.987)
53	80	0.9919	(0.987 - 0.995)
54	77	0.9524	(0.926 - 0.970)
55	70	0.9457	(0.914 - 0.966)
56	84	0.9557	(0.933 - 0.971)
57	85	0.9789	(0.968 - 0.986)
58	87	0.9733	(0.959 - 0.983)
60	80	0.9884	(0.982 - 0.993)
61	34	0.9674	(0.936 - 0.984)
63	91	0.9752	(0.963 - 0.984)
64	74	0.8453	(0.765 - 0.900)
65	62	0.9801	(0.967 - 0.988)

Relationship between test-retest reliability and data entry characteristics

There was no significant association between ICC and qualitative adherence category ($F_{1,24} = 0.12, p=0.76$). Linear regression of ICC (representing internal recording consistency) on the Adjusted Data Rate (measuring adherence to the recording

protocol) showed a weak and non-significant association(R-squared = 0.117, p = 0.086). These results are shown in Figure 7

Figure 7 Comparison of reliability of entry with adherence to diary schedule



Combining measures of adherence and reliability into a single measure.

The data in this section demonstrated heterogeneity of recording accuracy between subjects. Of the four measures used, three were measures of adherence to the diary

schedule (Adjusted Data Entry Rate, Paired Data Entry Rate and Qualitative Adherence Group) and overlapped strongly with each other, but not with the fourth (Intraclass Correlation Coefficient) which was a measure of the reliability, or reproducibility, of entries.

Both of these represent sources of error in the data, with some diaries being more error prone than others. One of the advantages of meta-analysis for subsequent comparison of between subject data, as used in later analysis, is that it allows for weighting of data according to its quality, giving greater influence to accurate data, and less to that which is more prone to error.

In order to generate a single weighting factor, a compound Adjustment Factor was derived for each participant as the product of the ICC and the Paired Data Entry Rate. The reason for choosing the paired data entry rate, was its wider range of values and its relevance to the later analysis of data pairs. The mean Adjustment Factor was 0.82, median 0.87 and range 0.57-0.99.

Reactivity

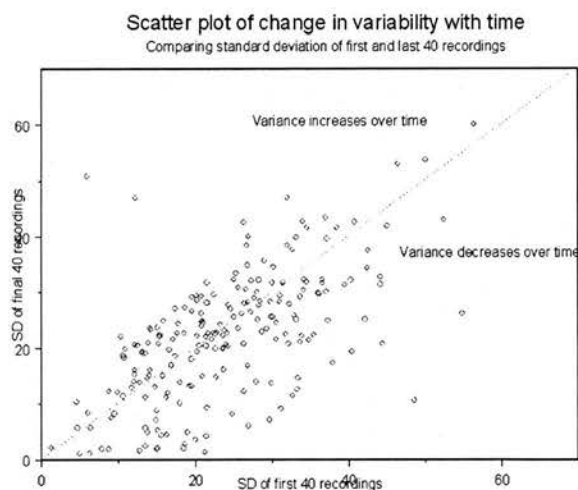
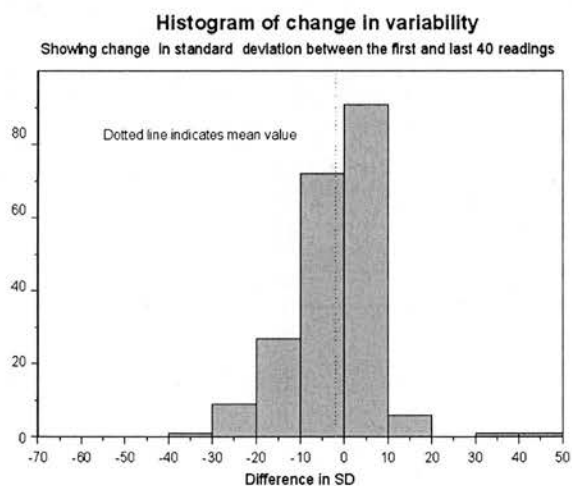
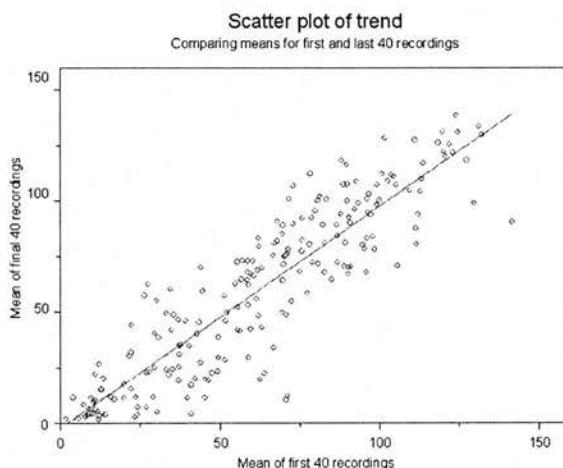
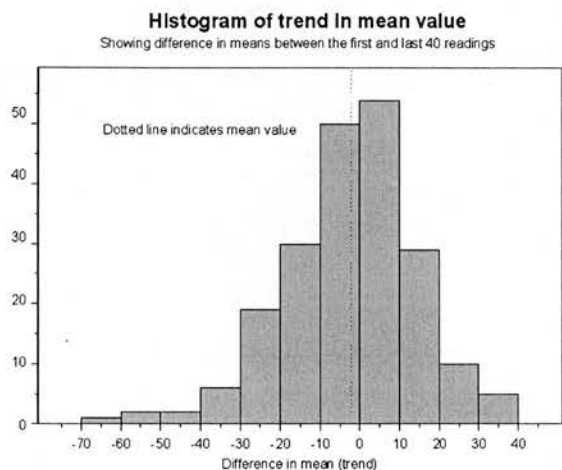
Reactivity is the unintended influence on the behaviour of a study participant simply through the process of research. Historically referred to as the Hawthorne effect this is likely to reflect both extra attention paid to the variable of interest and the feedback from participation(228). In the case of a diary study, the concern is that by recording self report measures the participant becomes more aware of their symptoms such that they may become amplified. Reactivity has not been a problem in shorter diary studies(158;161;167) but has not been reported in a longer study such as this one.

Reactivity was measured using only data from days on which two entries had been made in accordance with the diary schedule in order to avoid any bias from diurnal variation in symptoms. The first 40 and the last 40 readings in each series were used and where there were less than 80 readings (two participants) the samples were allowed to overlap rather than be shortened.

Although individual participants and variables showed some considerable variation, the pooled change was small. There was a small decrease in the pooled mean value

from 62.6 to 60.3. The mean standard deviation also reduced by a small amount, falling from 24.1 to 22.2. In the context of scales ranging from 0-150, these average changes are small. The distribution of changes is shown both as histograms of the changes in mean and SD, and scatter plots of the early and late values of mean and SD in Figure 8.

Figure 8 Tests for reactivity: changes in mean and standard deviation over time



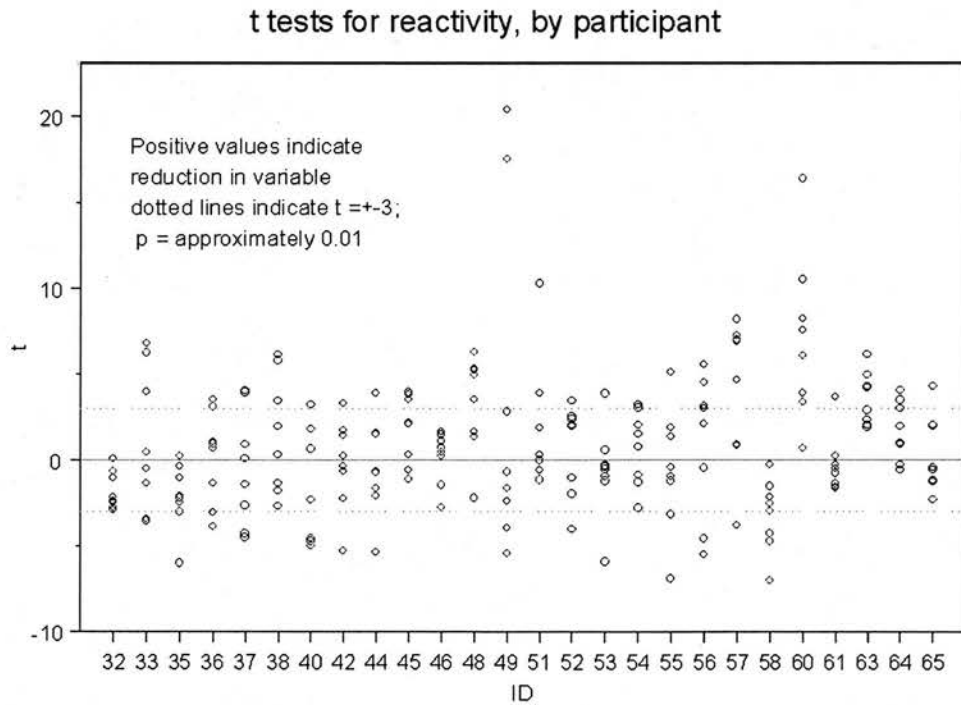
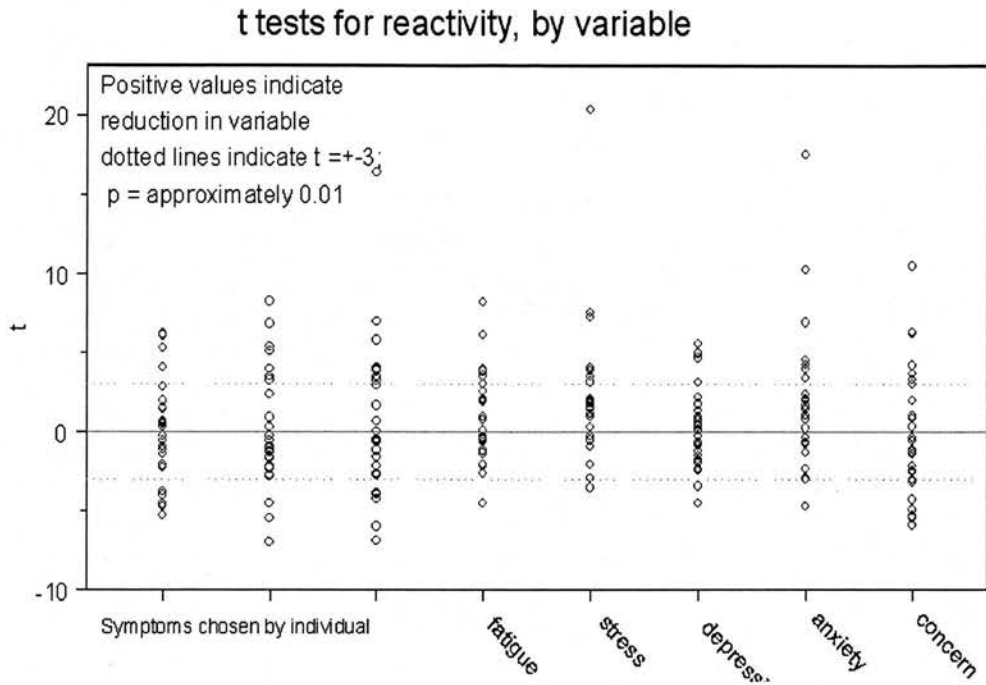
A statistical comparison of the changes was made using non-parametric techniques as not all of the changes were normally distributed. Wilcoxon signed rank test and showed the change in mean was not significant ($Z= 1.37, p= 0.17$), but that the change in SD was ($Z=2.31, p=0.02$).

There was significant heterogeneity between individuals in the value of change in the mean (Kruskall Wallis rank sum $\chi^2=56, df=25, p=0.0003$) but not between items ($\chi^2=10.3, df= 7, p=0.164$). For change in standard deviation the heterogeneity between individuals and items was highly significant ($p=0.007$ and $<.001$ respectively).

To look for patterns in changes at the level of the individual, t-tests were carried out for each individual and variable combination, between the first 40 and the final 40 data points. Figure 9 shows the distribution of t-test results, grouped by variable and by individual. The data for participants' chosen symptoms represent a mix of different symptoms and are simply grouped according to the order in which they were chosen by the individual. Because of the different numbers of participants recording each possible symptom, no attempt was made to compare individual physical symptoms for reactivity.

Comparison of t-test results between individuals shows broad similarities although two exceptions occur (subjects 49 and 60) in both of these there was more marked reactivity than in other participants. Four of the five outlying readings with t scores >10 represented stress or anxiety and all displayed a reduction in symptom rating as the study progressed.

Figure 9 tests for reactivity in mean value by individual and variable



Diurnal Variation in scores

The five common variables and the five most commonly chosen symptoms were analysed to assess the extent of diurnal variation. A simple arithmetical mean of the difference between the first and second readings for each variable was taken using only data from days with matching pairs of entries.

Table 13 shows the mean differences, with 95% confidence intervals and the result of a single sample t test for difference from zero. Fatigue increased over the day (mean difference 4.8, $p < 0.001$) but other changes were smaller and anxiety was the only psychological value to show a significant change although this was small (reduction of -1.3, $p = .034$). Among physical symptoms selected by participants, joint pain increased through the day whereas bowel symptoms reduced. Two entries are shown for bowel symptoms in Table 13: the first represents the mean difference for all participants reporting this. Further inspection of this difference, which is larger than the others, showed that one participant (44) had a much greater mean difference than the others. Hence the second result, which only achieves modest significance is calculated after exclusion of the outlying participant.

Table 13 Diurnal variation in symptoms:

mean difference represents the average change in value over the day from first to second data entry (negative values indicate reduction in score)

	Mean Difference	95% Confidence Interval		t	df	Sig. (2-tailed)
		Lower	Upper			
Fatigue	4.80	3.53	6.07	7.39	1,765	0.000
Stress	0.23	-0.99	1.44	0.37	1,765	0.714
Depression	0.88	-0.32	2.08	1.43	1,765	0.152
Anxiety	-1.30	-2.49	-0.10	-2.12	1,765	0.034
Concern	-0.55	-1.48	0.38	-1.16	1,765	0.245
Muscle pain	-0.74	-2.17	0.70	-1.01	1,072	0.314
Headache	-1.22	-3.31	0.87	-1.14	746	0.253
Joint pain	2.01	0.26	3.75	2.26	765	0.024
Bowel problems	-6.95	-9.08	-4.81	-6.39	653	0.000
Bowel problems*	-2.03	-3.83	-0.23	-2.21	571	0.028
Abdo Pain	1.36	-1.08	3.81	1.09	620	0.274

Conclusion: performance of electronic diaries

Overall the diaries showed good to excellent accuracy with good to excellent adherence to the data collection schedule. There were relatively few days lost to technical failure and the data was successfully extracted for analysis. A valid measure of data accuracy for each participant was developed for use in further analysis.

Testing for reactivity showed that there was a small, statistically significant but clinically unimportant, downward trend in both the mean and variance of symptom rating as the diary progressed.

Overall, the data was suitable for more detailed analysis. The results in this chapter confirm that electronic diaries are capable of generating data suitable for symptoms research.

Chapter 12 Results - preliminary time series analysis

Introduction

This section describes the preliminary time series analysis of data from the electronic diaries. It describes the process of selecting the most appropriate set of data and minimising the inherent biases of time series data. The purpose of this stage in the analysis is to optimise the subsequent tests of associations between symptoms and psychological variables.

In addressing variation and autocorrelation within the data the results in this section address research question 2: *how do symptoms and emotional states reported by electronic diary vary over time in patients with persistent MUS?* In particular it addresses the hypothesis that symptom diary data should show autocorrelation and that this can be dealt with statistically to permit further analysis.

Inspection of the data in the previous section showed that the overall data quality was good, and that while individual participants varied in their adherence to the data collection schedule and in their accuracy of recording, these could be quantified and adjusted for. As these data represented an interrupted time series, three questions required to be addressed before more detailed analysis could be carried.

- (1) which was the best form of the data to use?
- (2) how could the data be described in terms of serial dependence, trend and periodicity in order to select the best method for analysis?
- (3) how should the data be transformed to maximise the reliability of the later analysis?

Choice of data

In a few instances the data represented an almost completely intact time series, with no missing entries. For most participants however this was not the case and the series was incomplete. In the case of this study, with twice daily data entry, this could pose particular problems where diurnal variation in one or more variables was present. As all participants recorded a measure of fatigue, it was reasonable to anticipate that diurnal variation would be important and account for this. Thus, the set of daily

paired data (as described in the previous section) was chosen over a longer but more incomplete whole series for each participant.

Describing the nature of the time series data

Time series data can be described in terms of several factors: the overall trend, cyclical periodicity, dependence on previous values and the effect of other perturbations (including both the effects of variables under investigation and random variation). Before investigating the effect of other variables, it is necessary to reduce potential error due to other characteristics of the data.

Analysis of the temporal trends in data was reported in Chapter 11; to recap there was a small but statistically significant shift overall in the mean value but this trend over time was less than 5% of the original value for two thirds of participant-symptom pairs.

Periodicity was thought, a priori, to be unlikely in the data except in terms of diurnal variation. Menstrual cycles were considered as a possible factor, however the relatively short duration of the study (no more than three normal menstrual cycles) would make identification of this scale of periodicity unlikely. None was apparent on simple inspection of the data.

Testing for autocorrelation, the serial dependence of one value on its preceding value(s), is described below.

Results of autocorrelation function

The autocorrelation function (ac.f) for each participant-variable pair was estimated using a script in the R statistical system. Results were displayed in a spreadsheet and each ac.f was qualitatively graded into one of the following categories

Short autocorrelation: displaying a rapid decay in ac.f below the significance line before lag 3 (92 series)

Medium autocorrelation: with decay in ac.f before lag 5 (38 series)

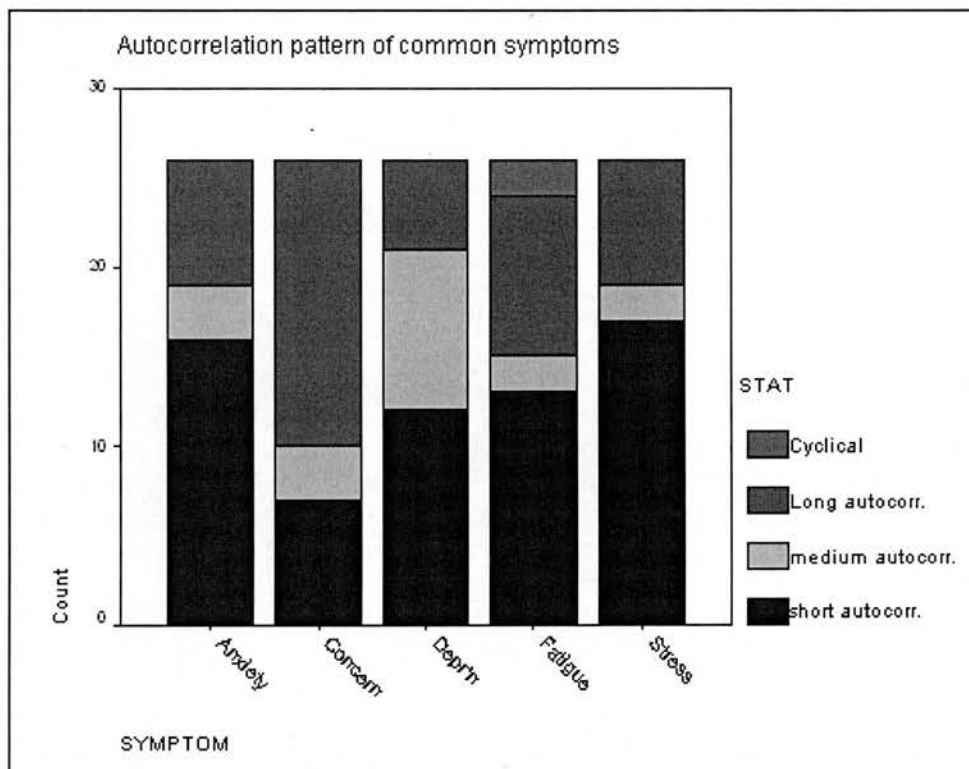
Long autocorrelation: with slow decay in ac.f which remained significant after lag 5 (74 series)

Periodic, oscillating between significant positive and negative values (4 series)

Of the four data series which showed clear oscillation, two were for fatigue (subjects 55 and 65) and one each for numbness (subject 53) and bowel symptoms (subject 44).

Comparison of the types of autocorrelation seen between the five symptoms common to all participants showed heterogeneity (Cramers V = .278, p= .003). Figure 10 shows the proportions of subjects in each autocorrelation category for the common variables and shows concern having the most individuals with long autocorrelation series and depression the least. This suggests that concern may have been the most stable variable and mood the least dependent on values more than 3 days earlier.

Figure 10 Autocorrelation category of common symptoms

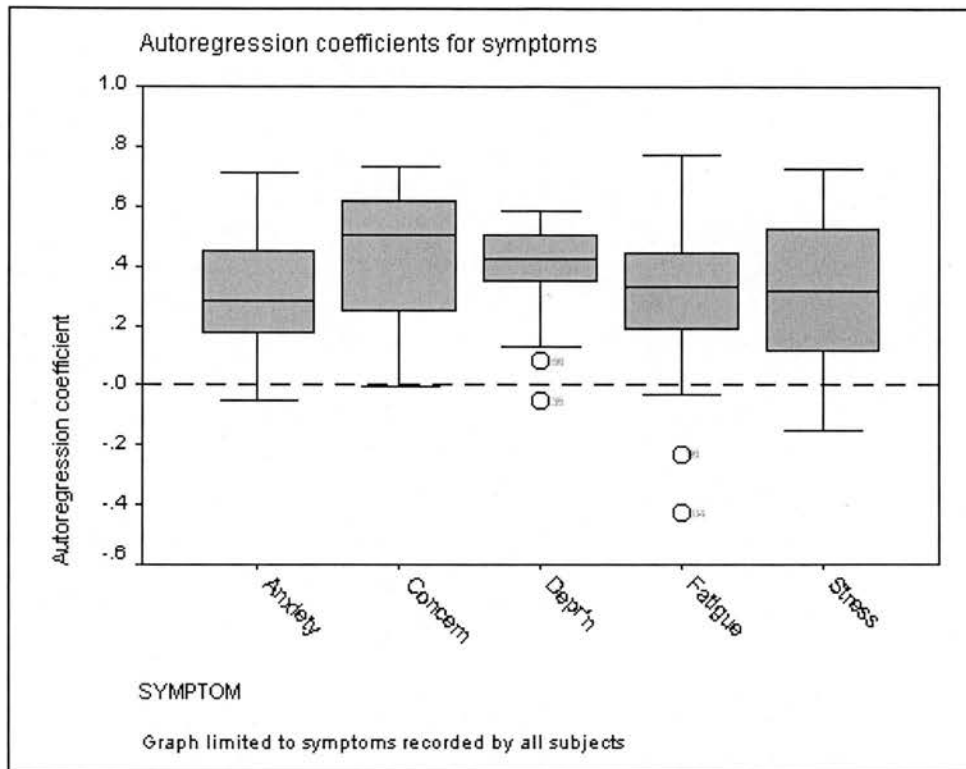


Differences between symptoms in measures of autoregression.

The autoregression coefficient for each participant-variable pair was calculated using a script in the R statistical system. The results for the five common variables are shown in Figure 11. As with the categorisation of autocorrelation functions in Figure 10, concern shows the highest coefficient, in keeping with greater dependency on past values, although using this approach, depression is not the lowest. Univariate

ANOVA suggested heterogeneity of borderline statistical significance ($F_{4,125}=2.51, p=.045$).

Figure 11 Box plot of autoregression coefficients



Effect of differencing the data.

Differencing of time series data (in which a series of actual values is converted to a series of differences between values) is a recognised method of removing autocorrelation(183), although it will not necessarily remove trend. Tests for reactivity as described earlier (which were designed to detect trends in mean and variance) and autocorrelation functions were repeated on the differenced data. Results were as follows.

Mean (first 40 points) = -0.03 (95% CI -0.172 to 0.09)

Mean (final 40 points) = -1.63 (95% CI -1.47 to -1.78)

Difference between means = -1.6, (Wilcoxon rank-sum test $Z= -13.1$ $p<0.0001$)

SD (first 40 points) = -28.65 (95% CI -26.8 to -30.5)

SD (final 40 points) = -25.59 (95% CI -23.7 to -27.5)

Difference between SDs = -3.1 (Wilcoxon rank-sum test $Z= 2.06$ $p=0.039$)

All autocorrelation functions for differenced data decayed below the significant level at lag(1) and t-tests for all autoregression coefficients were non-significant.

The effect of differencing the data was to remove autocorrelation and to reduce, but not eliminate changes with time in trend and in variance in recordings.

Testing for normality of distributions.

In order to use regression models for analysis, data should be, at least approximately, normally distributed. To test for this a Wilks-Shapiro test was carried out on each differenced time series. For each series the test statistic W , and the probability p that the data came from a normal distribution around the mean was calculated. Of 208 tests, the W value was less than 0.96, suggestive of non-normality in 90 (see Figure 12a).

One possible explanation for this non-normality from observation of simple time series plots of the data, was that in some data series there was very little variation about the mean. This was particularly the case for the psychological variables such as anxiety, stress and concern about symptoms.

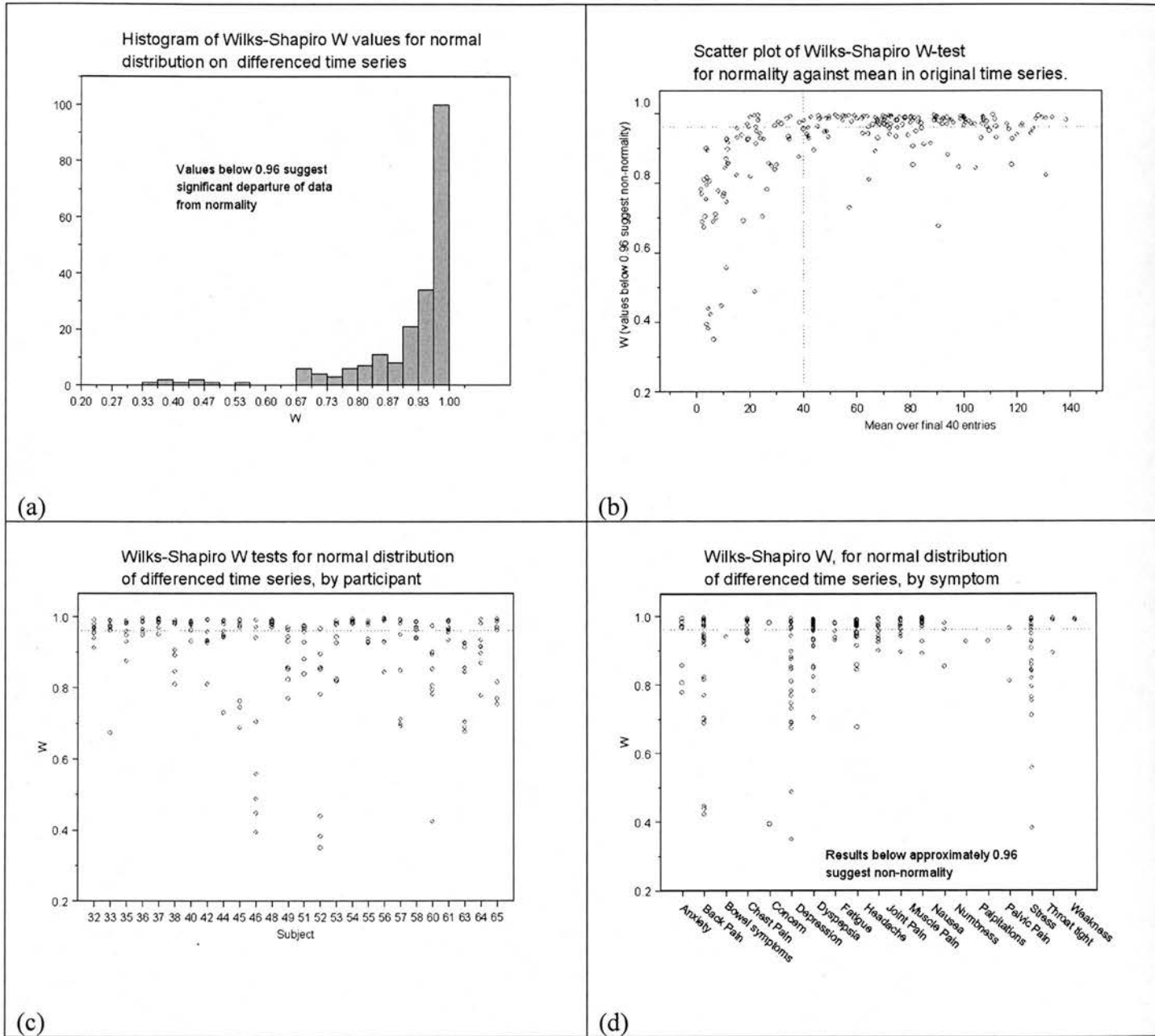
To test the possibility that the very low values for W occurred in series with very low variance, a plot was made of W against the standard deviation of the final 40 readings as used in the assessment of reactivity (

Figure 12b). This showed a predominance of low values for W among the time series with low mean values. Referring back to the original time series plots of data, these represent cases where one or more variable has been consistently rated at low levels over a sustained period of the diary data collection. These cases have low variability and generally did not feature in important correlations between variables, as will be shown later.

In order to explore this further, the W statistic was plotted by symptom and by individual participant (Figure 12 c and d). Visual inspection suggested that the psychosocial symptoms were associated with lower values of W (i.e. less likely to come from a normal distribution) and this was confirmed by a two sample test

between psychosocial and physical symptoms: Wilcoxon rank-sum test $Z=-4.25$
 $p<0.0001$.

Figure 12 Tests for normality in differenced time series



Conclusion: preliminary time series analysis

This section began with three questions concerning the data and these have been answered.

The daily pairs dataset appears robust and has advantages through preserving the diurnal time structure.

There are significant trends and autocorrelations, but no periodicity other than the diurnal one, which is relatively weak in most series.

The trends and autocorrelation can be removed by generating a differenced time series of the paired same day data to produce series which show only small trends in mean value and variance, which have minimal autocorrelation and which are, in most instances normally distributed. Those series which are not normally distributed are characterised by low mean and low variance.

While not meeting strict criteria for stationarity (183) and formal time series analysis, the data are adequate for further analysis using conventional correlation and regression statistics.

Chapter 13 Results - correlation within subjects.

Introduction

This section describes the analysis of correlation between symptom variables and in doing so addresses research question 3, *What are the concurrent associations between symptoms and mood?*, and tests the two hypotheses related to this: that there will be significant associations between physical and emotional variables and that these associations will differ within and between individuals.

The chapter begins by describing the correlation coefficients between variables at the level of the individual subject and then shows the results of pooling the individual findings using a random effects meta-analysis. This pooled analysis is then used to test for the effect of baseline measures of anxiety, depression, illness worry and tendency to experience symptoms on the correlation between variables. This analysis is confined to concurrent associations, using differenced data to avoid autocorrelation, and so its results cannot be taken to imply causation, but simply correlation of two variables which may be due to a direct relationship between them, but may equally be due to the presence of one or more confounding variables.

Generation of correlation matrices.

Correlation was calculated between each pair of symptoms in all 26 subjects. This was carried out using the Pearson product moment correlation coefficient r on the differenced series of daily paired data. As was shown in Chapter 12, while most series met the assumptions of multivariate normality necessary for the Pearson correlation, some did not. A sample of those was analysed to see whether the correlation coefficients from these non-normal variables was noticeably different from the normally distributed ones and there was no difference. In light of this, the risk of over-estimating correlation of a few data series through using parametric statistics was chosen over reducing the power across all series by using a non-parametric technique.

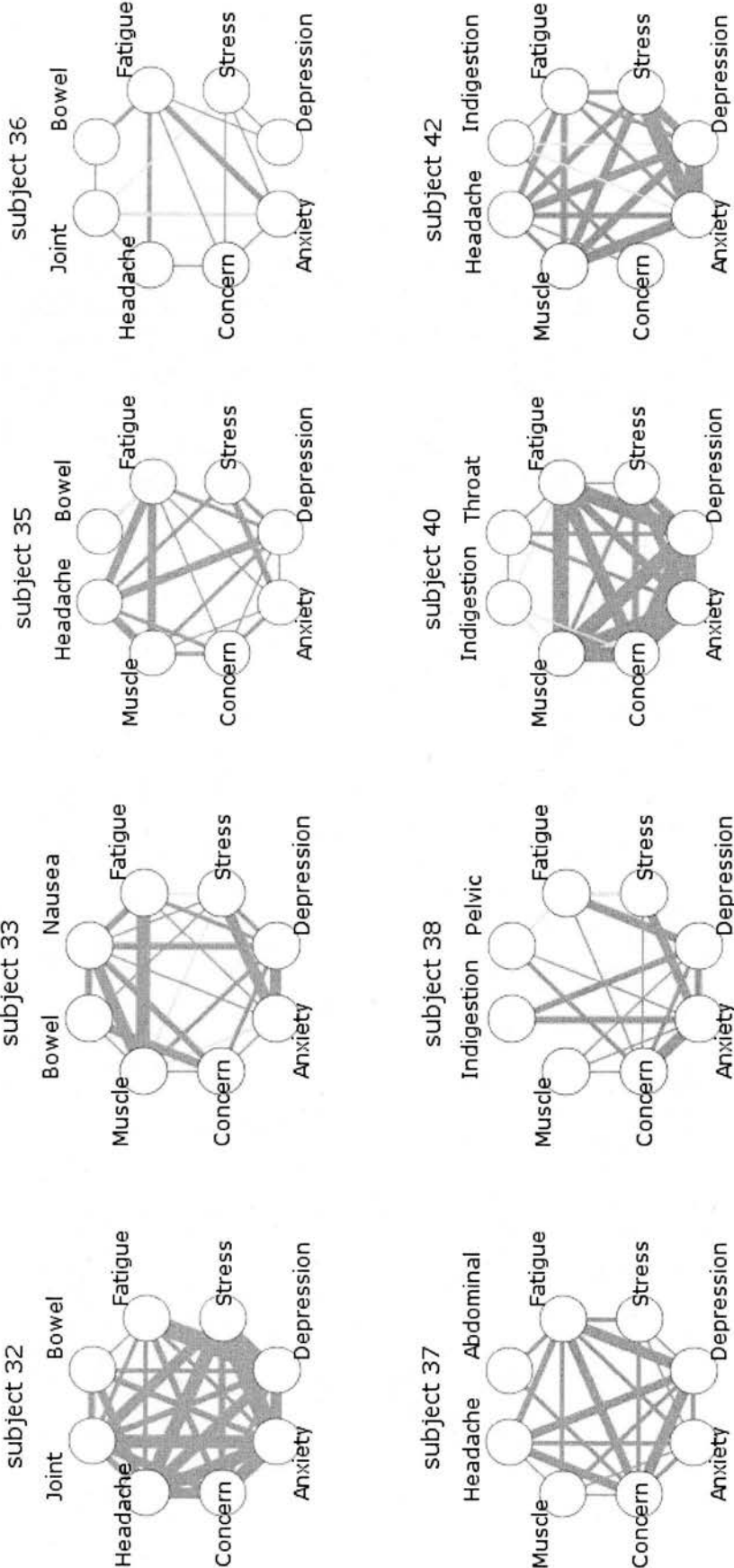
A correlation matrix was generated for each individual participant and these were stored in one master table and represented graphically as network diagrams in Figure 13. Visual inspection of Figure 13 shows there was considerable heterogeneity in the

number and strength of correlations between individuals, with some individuals' matrices strongly interconnected (e.g. subjects 32, 48,54) and others quite sparsely connected (36,46,52). In particular, there were also differences in the patterns and strength of correlations between physical symptoms and psychological variables, as evidenced by the links between the upper and lower halves of each matrix. Some individuals showed their predominant correlations with mood (37,51) while others showed stronger associations through stress or concern (45,54,60):

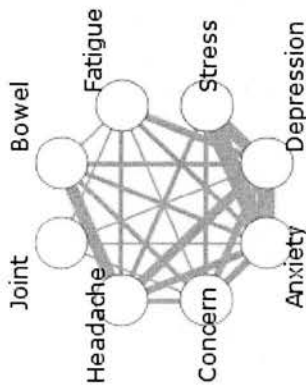
As examples of the scales used in the network diagrams, the correlation for subject 33 were as follows: broad line (Muscle Pain – Fatigue), 0.57; moderate line (Concern –Nausea), 0.32; thin line (Bowel-Stress) 0.08. A pale line represents negative correlations (Muscle –Stress),-0.08. Few correlations were negative, and all of these lay between zero and -0.15.

While 95% confidence intervals depended to some extent on sample size, the majority of individual subject correlation coefficients between -0.15 and 0.15 were not statistically significant.

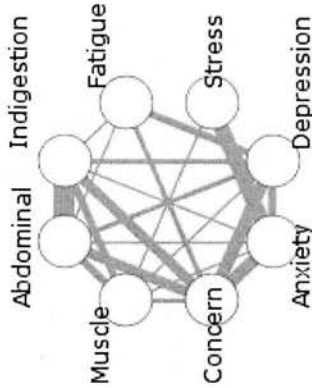
Figure 13. Network diagram of each individual subject correlation matrix. Line thickness of lines represent the size of r the Pearson correlation coefficient



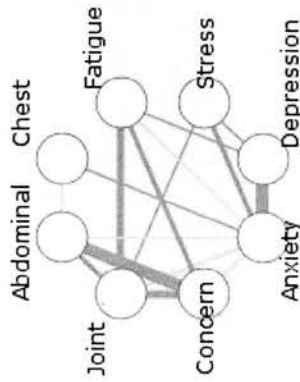
subject 44



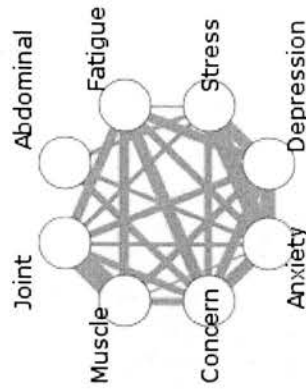
subject 45



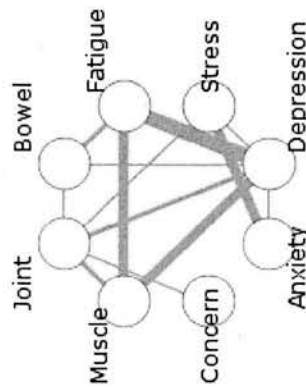
subject 46



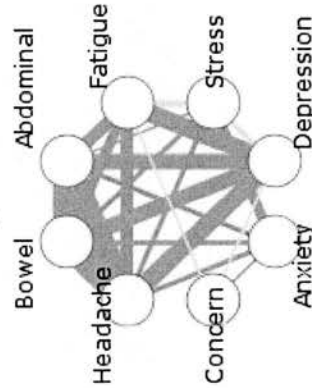
subject 48



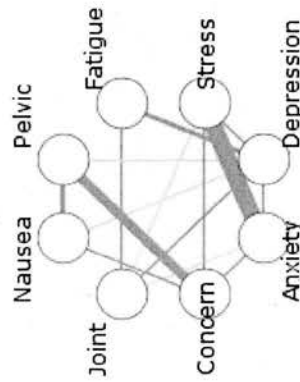
subject 49



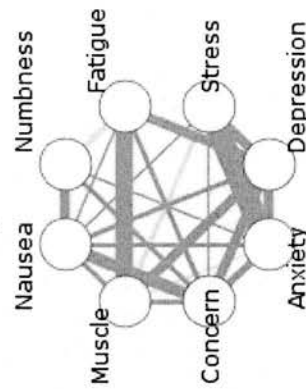
subject 51



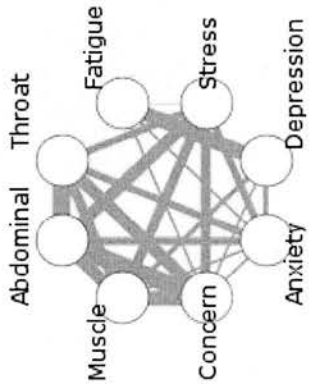
subject 52



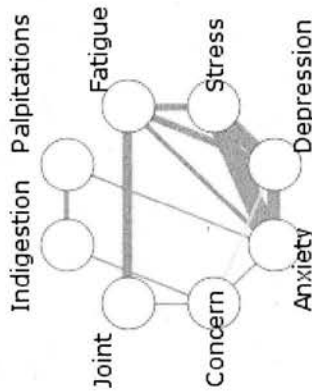
subject 53



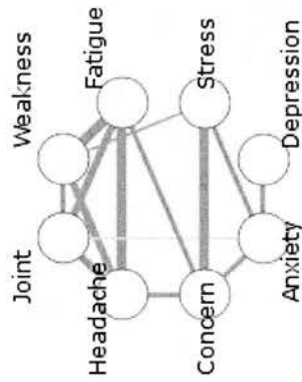
subject 54



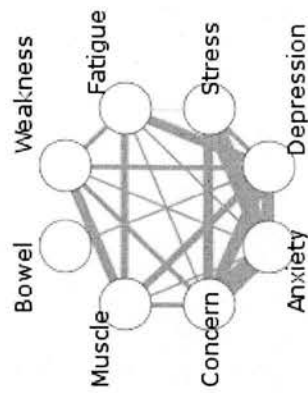
subject 55



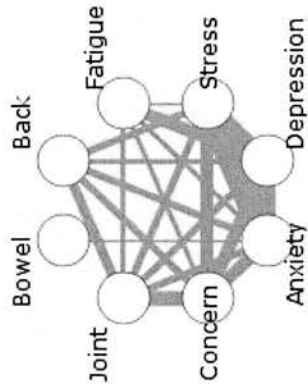
subject 56



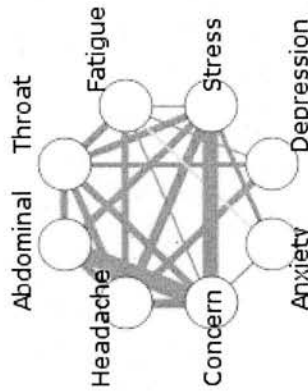
subject 57



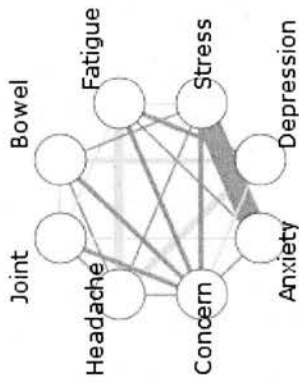
subject 58



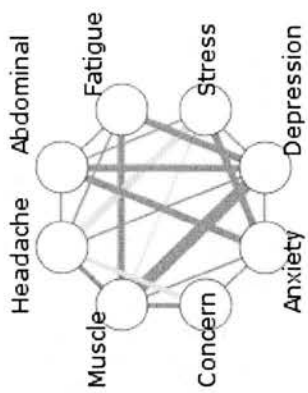
subject 60



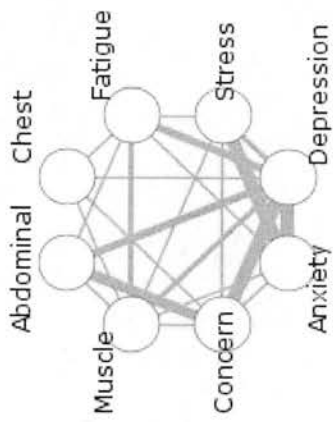
subject 61



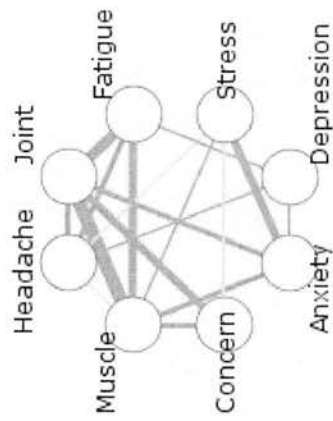
subject 63



subject 64



subject 65



Pooling of correlation coefficients by meta-analysis of regression coefficients

In order to make formal quantitative inter-individual comparisons a formal meta-analysis was carried out. Because each participant had collected data on their choice of 3 personally relevant symptoms from a list of 14, each individual's matrix had their own choice of variables. Therefore rather than comparing data together, pooling and meta-analysis was carried out for each variable pair (e.g. concern and fatigue). The number of matching correlation coefficients ranged from 26 (for combinations of the common variables: fatigue, depression, anxiety, stress and concern) to one. Meta-analysis was only carried out on variable pairs which occurred in at least three participants

Because of the need to carry this out on a large number of sets of correlation coefficients a computer script was written in Python following the method for meta-analysis of correlation coefficients detailed by Hunter & Schmidt(222). This differs from that published by Field(188) only in the use of confidence intervals derived directly from the variance in the data (using the method outlined by Whitener(224)) and in the presentation of credibility intervals in addition to confidence intervals.

The steps of the meta-analysis were as follows:

- Calculation of confidence interval for each coefficient r
- Adjusting of each individual r for error in data collection. This was carried out using an adjustment factor to account for error in adherence and reliability as described in Chapter 11
- Weighting the value of individuals' r by sample size
- Calculating the meta-analytic value for r based on the weighted adjusted coefficients
- Estimation of confidence intervals and credibility intervals for meta-analytic coefficients.

The effect of adjusting for the measurable error in data collection and weighting according to sample size, was to increase the value of the corrected correlation

coefficients in those data sets with greater error (i.e. lower adjustment factor) and to give greater prominence to coefficients from larger samples.

The results of the multiple meta-analyses were presented in two forms, a matrix containing the results of meta-analysis, and as a set of forest plots for each variable pair. The matrix is shown in Table 14.

Table 14 Meta-analytic correlation coefficients for variable pairs

	Fatigue	Stress	Depr'n	Anxiety	Conc'n	Musc	Head	Joint	Abdo
Fatigue									
Stress	0.02								
Depression	0.41	0.26							
Anxiety	0.13	0.59	0.42						
Concern	0.20	0.23	0.26	0.32					
Muscle pain	0.40	0.09	0.36	0.17	0.30				
Headache	0.35	0.15	0.31	0.18	0.21	0.25			
Joint pain	0.24	0.08	0.15	0.05	0.25	0.57	0.24		
Bowel probs	0.08	0.04	0.10	0.09	0.11	0.03	0.43	0.08	
Abdo pain	0.11	0.20	0.17	0.18	0.46	0.21	0.29	0.13	
Dyspepsia	0.02	0.02	0.08	0.08	0.21	0.09			
Nausea	0.15	0.06	0.13	0.14	0.36	0.32			
Tight throat	0.18	0.26	0.14	0.22	0.33	0.26			0.50
Chest pain	0.10	-0.01	0.16	0.12	0.17				-0.07
Pelvic pain	0.03	0.04	-0.14	0.06	0.43				
Weakness	0.31	0.08	0.17	0.10	0.16				

Differenced(1) data. Meta-analytic correlation coefficients with 95% CI excluding zero are shown in bold

The strongest correlations appeared between muscle and joint pain ($r = 0.57$, number of subjects. $N=5$) and stress and anxiety ($r = 0.59$, $N=26$). Both are likely to have represented overlap between two similar constructs which were difficult to differentiate. Depression was moderately correlated with anxiety ($r=0.42$), but also with fatigue ($r=0.41$), muscular pain ($r=0.36$) and headache ($r=0.31$). Fatigue was moderately correlated with depression ($r=0.41$), muscle pain ($r=0.40$), weakness ($r=0.31$) and headache ($r=0.35$). While stress and anxiety were only weakly correlated with physical symptoms (≤ 0.2 for all except throat tightness), concern

about symptoms was significantly correlated with all physical symptoms including abdominal pain ($r=0.46$), pelvic pain ($r= 0.43$), nausea ($r=0.36$), throat tightness ($r=0.33$) and muscle pain ($r=0.30$). As the participants all had unexplained symptoms and were generally seeking explanation for them, it is important to recognise that the correlation of concern and symptoms would be anticipated, and does not imply a causal relationship in either or both directions.

The apparently strong correlation between abdominal pain and throat tightness reflects a very strong correlation ($r=0.8$) in one of the three participants recording these two items.

For each variable pair the distribution of individual correlation coefficients was displayed as a forest plot of each meta-analysis. In addition to the mean corrected correlation coefficient for each variable pair, the plots included confidence intervals for each individual's correlation coefficients and for the mean corrected coefficient, and credibility intervals as a qualitative measure of heterogeneity of the individual studies. In almost all plots with more than 5 participants the credibility intervals were wide (either greater than the 95% confidence intervals, including zero or both of these) suggesting heterogeneity of the elements in the meta-analysis. shows a sample of the plots.

Figure 14 - Examples of forest plots of meta-analyses of correlation coefficients.

Figure 14a Examples of meta-analysis of correlations. Dotted lines indicate pooled r and credibility intervals

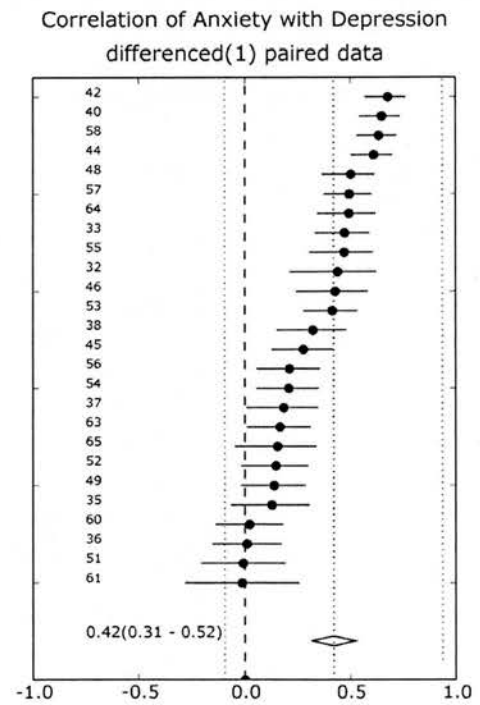
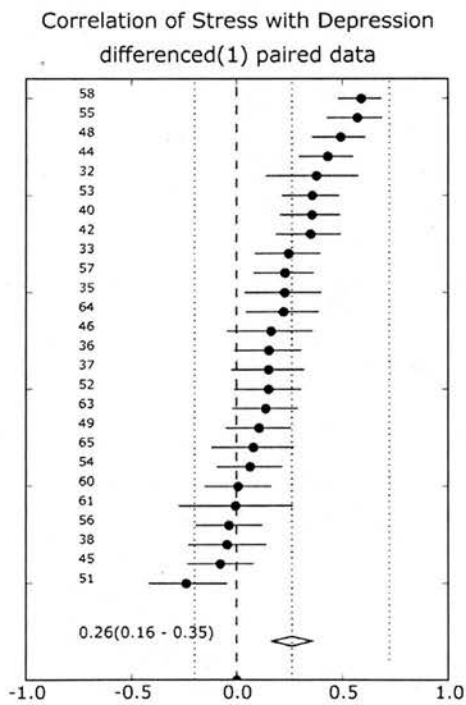
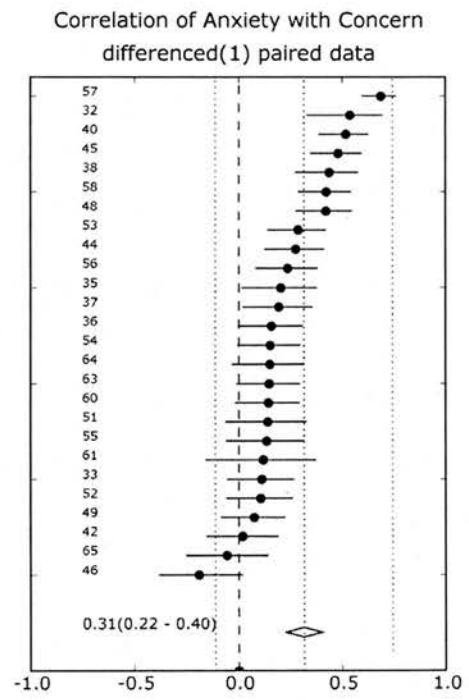
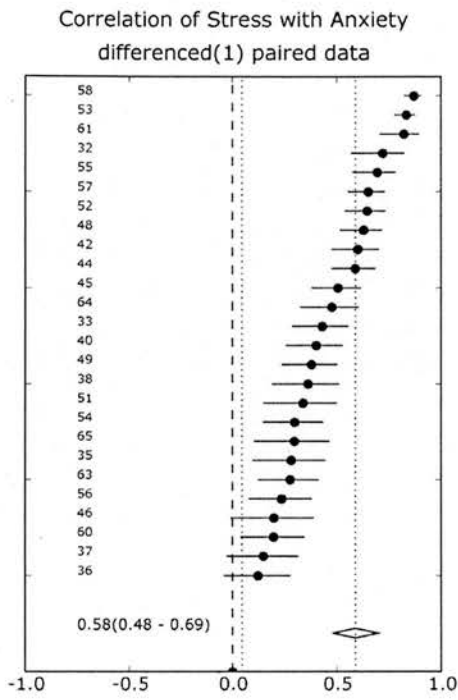
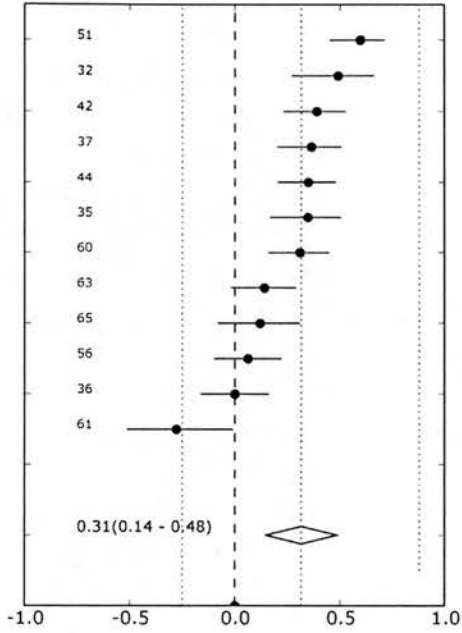
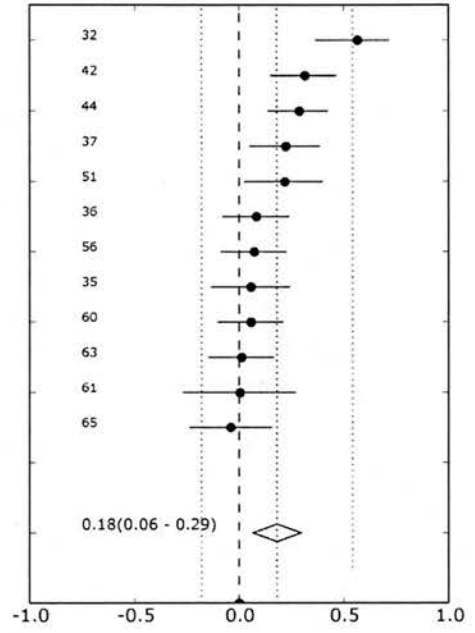


Figure 14b Examples of meta-analysis of correlations. Dotted lines indicate pooled r and credibility intervals

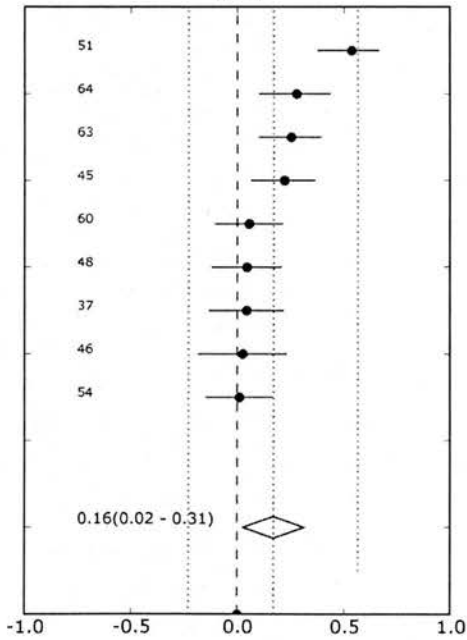
Correlation of Depression with Headache
differenced(1) paired data



Correlation of Anxiety with Headache
differenced(1) paired data



Correlation of Depression with Abdo pain
differenced(1) paired data



Correlation of Concern with Abdo pain
differenced(1) paired data

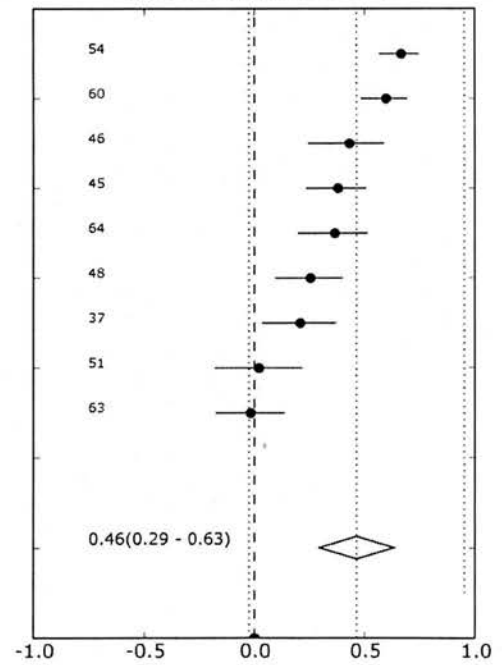
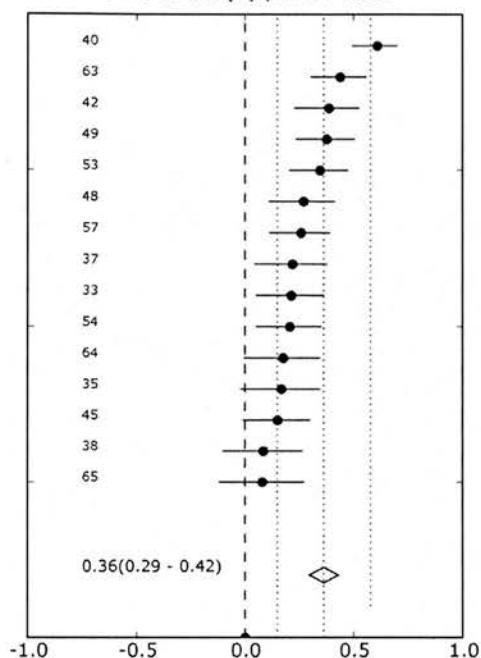
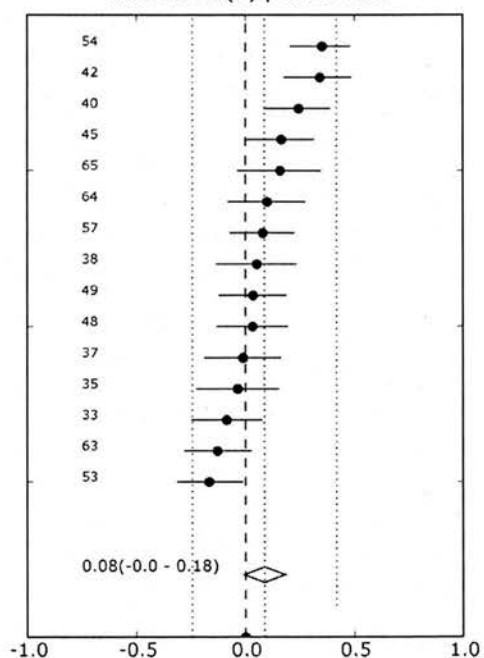


Figure 14c Examples of meta-analysis of correlations. Dotted lines indicate pooled r and credibility intervals

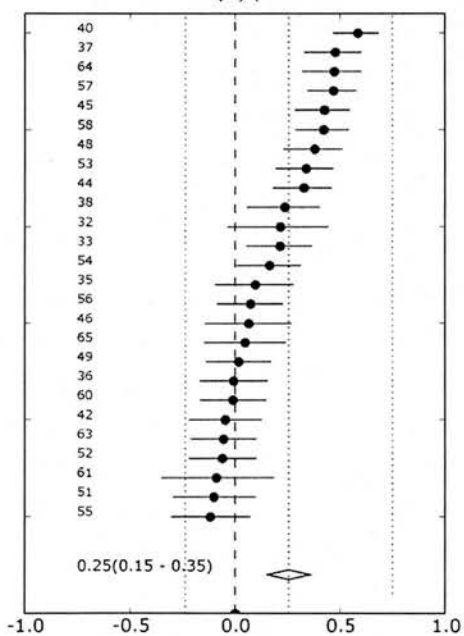
Correlation of Depression with Muscle Pain
differenced(1) paired data



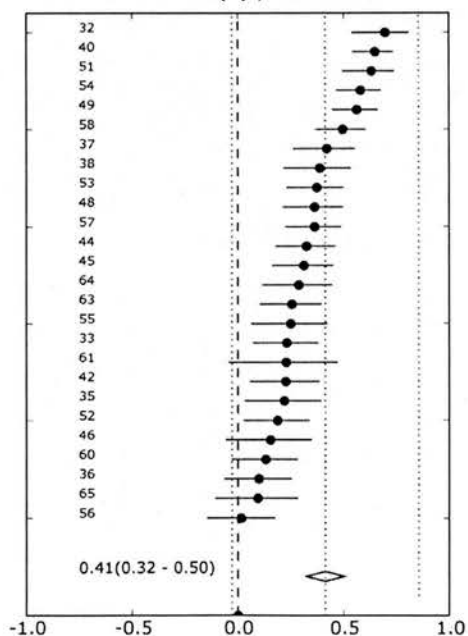
Correlation of Stress with Muscle Pain
differenced(1) paired data



Correlation of Depression with Concern
differenced(1) paired data



Correlation of Fatigue with Depression
differenced(1) paired data



Heterogeneity of correlation coefficients

While there were many significant correlations between variable pairs in the meta-analysis there was also considerable heterogeneity between individuals. This was demonstrated in the wide credibility intervals and the number of individuals for whom the 95% confidence intervals for correlation coefficients did not include the meta-analytic r . An exploratory analysis was thus made using baseline questionnaire data to address two additional questions: does the baseline level of anxiety, depression, illness concern or somatisation influence the strength of correlations between variables, and if it does, is the effect limited to correlations involving that factor?

For the four scales, HAD-Anxiety, HAD-depression, Illness Concern (Whitely 7) and Somatic Symptoms Inventory, the sample was split at the median into two approximately equally sized groups: the high and low scorers for that scale. The meta-analyses were all repeated separately for high and low scorers on each scale and the results were compared by paired value t-tests.

Higher between symptom correlations were seen in those subjects who scored above-median baseline levels on the HAD-Anxiety ($t=4.82$, $p<0.001$); HAD-Depression ($t=2.96$, $p=0.005$) and Illness Concern Questionnaire ($t=4.59$, $p<0.001$) scales. The difference was smaller and non-significant for the Somatic Symptoms Inventory. To test whether this was simply due to the baseline variable affecting correlations including itself (for instance HAD-Anxiety being associated with higher correlations involving anxiety) tests were repeated for correlations excluding those related to the baseline measure. The results of this are shown in Table 15.

Table 15 Comparison of meta-analytic correlation coefficients between high and low scores on baseline scales.

Baseline; scale	Comparison includes	t	Df	Significance
Anxiety; HAD A	All correlations	4.82	36	<.001
	Only including anxiety	3.38	8	.010
	Only excluding anxiety	3.67	27	.001
Depression; HAD D	All correlations	2.96	42	.005
	Only including depression	1.73	9	.118
	Only excluding depression	2.37	32	.024
Concern; ICQ	All correlations	4.59	40	<.001
	Only including concern	1.95	8	.086
	Only excluding concern	4.26	31	<.001
Somatisation; SSI	All correlations	1.17	45	.246

To investigate whether some correlations were more sensitive to baseline variables than others, the variable pairs with the greatest difference between high and low baseline scorers were identified. Table 16 shows those meta-analytic correlation coefficients with the greatest difference between high and low baseline scorers.

Magnitude of the correlation coefficients

The calculation of correlation coefficients described here quantifies the interaction of separate variables and shows that many of these are statistically significant. However this is not an easy way of quantifying their importance, either to a patient with unexplained symptoms or to a doctor helping them make sense of their experience. One useful interpretation is that r^2 , the square of the Pearson correlation coefficient is a measure of the reduction in unexplained variance in one variable through knowing the other. Thus, typical values seen in this study for the correlation coefficient r of 0.5, 0.3 and 0.1 respectively represent a 25%, 9% and 1% reduction in variance which is equivalent to saying that the two variables account for 25%, 9% and 1% respectively of the variation in each other.

In absolute terms, then, the influence of variables seen in the correlations from the diary data are relatively modest. Even choosing a fairly strong association such as between low mood and headache, with pooled correlation coefficient 0.31 (95%CI

0.14-0.48), this translates to accounting for approximately 10% (95% CI: 2%-23%) of the variance.

Table 16 Meta-analytic correlation coefficients showing the largest differences between high and low scorers on baseline scales

Scale		Low Scorers		High Scorers	
Correlated variables		R	(95% CI)	r	(95% CI)
HAD - Anxiety					
Depression	Headache	0.19	(-0.05 - 0.42)	0.53	(0.41 - 0.64)
Concern	Abdo Pain	0.40	(0.21 - 0.59)	0.69	(0.57 - 0.80)
Anxiety	Headache	0.08	(0.00 - 0.15)	*	0.36 (0.17 - 0.54)
Anxiety	Joint Pain	-0.09	(-0.22 - 0.03)		0.18 (0.01 - 0.35)
Fatigue	Depression	0.30	(0.19 - 0.41)	*	0.55 (0.44 - 0.66)
HAD -Depression					
Depression	Headache	0.23	(0.05 - 0.42)	*	0.66 (0.58 - 0.74)
Concern	Dyspepsia	*	0.43 (0.32 - 0.54)		0.04 (-0.16 - 0.23)
Headache	Bowel	*	0.48 (-0.07 - 1.04)		0.14 (-0.50 - 0.78)
Anxiety	Headache	0.12	(0.04 - 0.21)	*	0.44 (0.19 - 0.70)
Stress	Joint Pain	-0.06	(-0.13 - 0.01)	*	0.25 (0.11 - 0.40)
Concern	Headache	0.15	(0.01 - 0.30)		0.45 (0.20 - 0.70)
Anxiety	Joint Pain	-0.07	(-0.18 - 0.03)		0.21 (0.00 - 0.43)
ICQ					
Headache	Bowel	0.01	(-0.11 - 0.12)	*	0.78 (0.14 - 1.41)
Headache	Abdo Pain	0.22	(0.10 - 0.33)		0.69 (-0.82 - 2.19)
Headache	Depression	0.18	(0.04 - 0.32)	*	0.61 (0.35 - 0.86)
Headache	Anxiety	0.06	(-0.01 - 0.13)		0.46 (0.30 - 0.61)
Muscle	Abdo Pain	0.01	(-0.11 - 0.14)		0.29 (0.07 - 0.51)
Fatigue	Depression	0.28	(0.21 - 0.36)	*	0.55 (0.42 - 0.68)
SSI					
Concern	Dyspepsia	-0.05	(-0.16 - 0.07)	*	0.39 (0.29 - 0.50)
Concern	Abdo Pain	*	0.63 (0.53 - 0.74)		0.27 (0.05 - 0.48)
Muscle	Dyspepsia	-0.11	(-0.24 - 0.01)		0.20 (-0.04 - 0.45)
Abdo Pain	Stress	*	0.33 (0.20 - 0.45)		0.04 (-0.04 - 0.12)
Fatigue	Dyspepsia	-0.15	(-0.25 - -0.04)	*	0.13 (0.01 - 0.25)
Depression	Abdo Pain	0.04	(-0.05 - 0.13)		0.32 (0.09 - 0.54)
* indicates non-overlap of confidence intervals, with marked column indicating the higher correlation					

Conclusion

This comparative analysis of correlation coefficients between multivariate time series data has not been reported previously. Analysis of correlation coefficients between symptoms, after removing the effect of autocorrelation by using differenced data, clearly shows statistically significant associations between symptoms at the same point in time.

Low mood and illness concern were both associated with more significant mind-body correlations than either anxiety or stress. In the case of low mood these were pain symptoms (muscular pain and headache) and fatigue which were all recorded by at least half of the subjects. Concern was significantly correlated with all symptoms but had strongest correlations with visceral symptoms such as pelvic and abdominal pain, nausea, throat tightness.

The meta-analysis of correlations showed substantial heterogeneity between individuals and this was partly accounted for by baseline scores for anxiety, depression and illness concern all of which showed greater correlation coefficients in high scorers at baseline. Somatisation, as measured by the somatic symptoms inventory, was not associated with differences in correlation between variables.

Statistically significant meta-analytic correlations accounted for a relatively small proportion of the total variance of the data, ranging from around 1% to 25%.

These correlations cannot be taken to imply causality and while demonstrating that physical symptoms are more highly correlated with mood and concern than anxiety or stress, cannot be used to argue that one caused the other rather than vice versa.

Chapter 14 Results - temporal sequences

Introduction

Thus far, the analysis has sought to break down each multivariate time series of patient data into isolated sets of data at each data collection time, with as little association between temporal neighbours as possible. The purpose of this, as demonstrated in Chapter 13, has been to measure correlations between variables at discrete points in time, while minimising within variable correlation.

This and the next chapter examine effects over time, in relation to research question 4: *Is there any evidence for consistent sequential relationships between symptoms over time?* This chapter will address the first hypothesis in relation to sequential relationships, that *time series data from patients with MUS will not show predictable sequential changes to suggest psychosomatic causal sequences* and Chapter 15 will test the hypothesis that *time series data will show an increase in concern and / or symptoms prior to GP consultation, in keeping with a previously published model.*

This chapter examines the extent to which variables in parallel series are dependent on prior values, both of themselves and of each other. To do this it uses a form of testing first developed in economics to use the preceding (lagged) values of a second variable to increase the prediction of the first variable. The technique is known as granger causality(184) after its originator, a Nobel Prize winning economist.

Granger Causality

Granger causality is deemed to be present when the addition of the lagged values of a time series X improves the prediction (i.e. reduces the unexplained variance) in the regression equation for series Y based on the lagged values of Y.

In view of the structure of the data with two data entries per day, and the lack of autocorrelation beyond lag 2 for the majority of (but not all) series, the statistical model was specified with lag(1) and lag(2) data. Repeating this with lag(1-4) data produced results which were only minimally different.

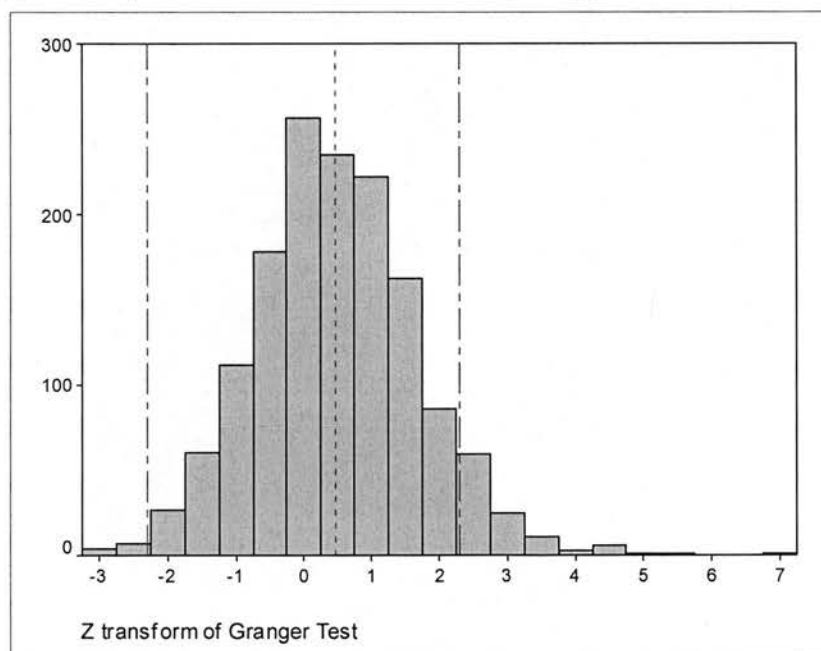
For each pair of variables, within-person regression models were calculated using a script written around the R `granger.test` module(221). The p value of each

comparison was transformed into a Z-score to generate a more easily manipulated output. (For reference, a p value of 0.025 is equivalent to z score of 1.96; p of 0.01, to z of 2.3; and p of 0.001 to z of approximately 3). The analysis was carried out on the paired daily data.

Results of Granger Causality test

1456 subject by variable pair combinations were tested, representing 26 subjects each with all possible combinations of the 8 variables to ensure that both directions of causality were tested. The Z scores of the tests were approximately normally distributed with a mean value of 0.48 and SD of 1.2 as shown in Figure 15. This can be interpreted as indicating that while most individual granger tests were not statistically significant ($Z < 2.3$), there was a net increase in the predictive value of the regression models by adding in lagged values of each independent variable.

Figure 15 Histogram of Z scores from Granger tests on all variable pairs (N=1456; dashed line indicates mean (0.48); dash-dot lines indicate $Z = \pm 2.3$)

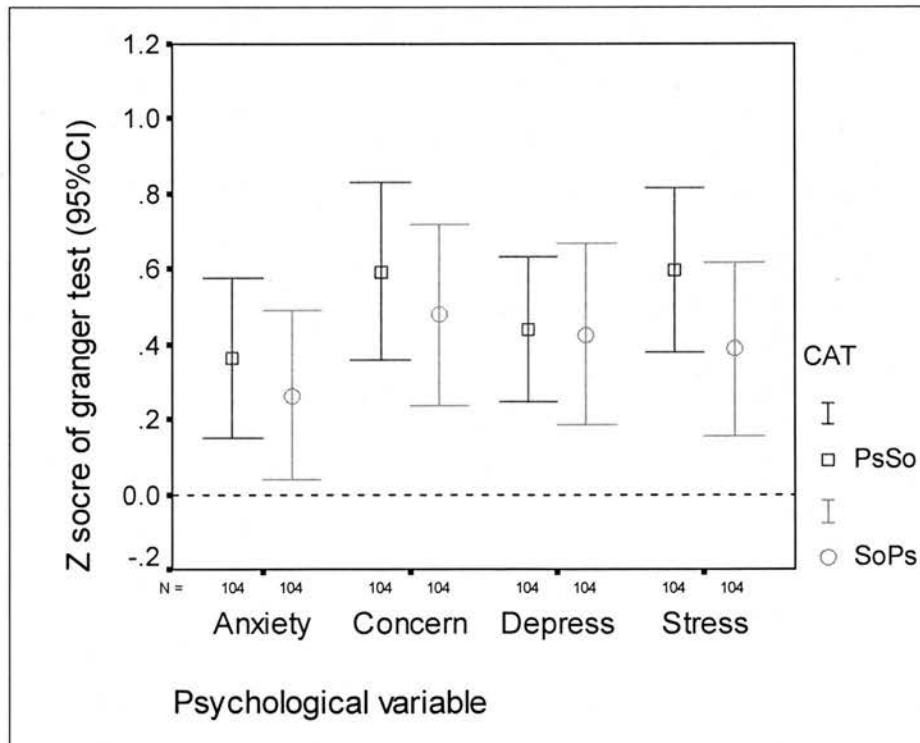


105 (7.2%) of granger tests yielded a p value < 0.01 (Z score > 2.3) whereas only about 15 (1%) would have been expected by chance. Results were categorised according to whether the interacting variables were somatic or psychological and listed such that the first-named “granger caused” the second. For example, in the somatic-psychological category, the lagged values of the somatic symptom

significantly increased the predictive accuracy of the model for the psychological one. Comparison of the number of significant results in each of the four categories showed no differences between categories: somatic-psychological (30 instances), somatic-somatic (25), psychological-somatic(25) and psychological-psychological(25).

In order to further investigate mind-body associations, results were broken down by individual symptoms and variables and the two directions of causation – psycho-somatic (PsSo) and somato-psychic (SoPs) – compared against each other. Figure 16 shows this for the four psychological variables, indicating that anxiety, stress and concern all have slightly greater effects in a PsSo than SoPs direction although these are not statistically significant either individually or for pooled data. Depression shows no difference between PsSo and SoPs causality. For each variable the lower 95% confidence interval is above zero in both directions of causality indicating some statistically significant directional effect.

Figure 16 Directional Granger Causality - Psychological Variables
PsSo = Psycho-somatic causation, SoPs = somato-psychic

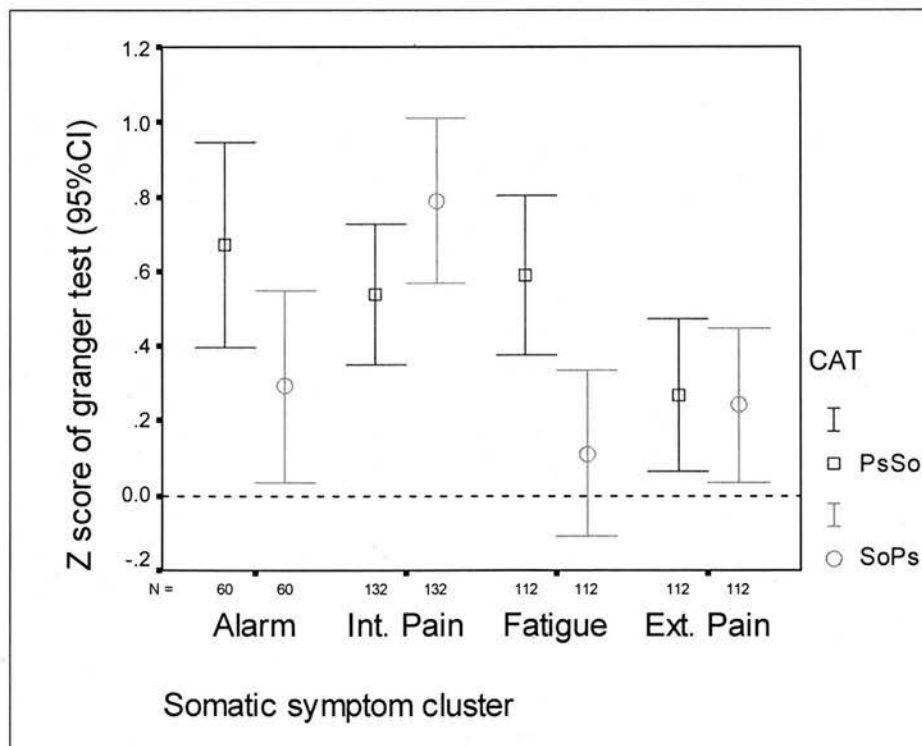


Because there was a larger number of physical symptoms to test, these were categorised into four clusters: alarm, internal pain, fatigue and external pain. The

alarm group included nausea, dyspepsia, swallowing difficulty, chest pain, numbness and palpitations; internal pain comprised headache, abdominal pain, pelvic pain and “bowel problems”; the fatigue cluster included fatigue and weakness, and the external pain included muscle, joint and back pain. This grouping was derived by combining physiological features of the symptom with similarities in granger tests: i.e. it was a post hoc classification.

The granger test results for these categories are shown in Figure 17. All results except the SoPs direction for fatigue have 95% confidence intervals above zero. The four groups appeared different. The alarm group showed significantly higher PsSo than SoPs influence ($F_{(1,118)}=4.09$ $p=0.045$); the internal pain group showed a reversed trend with higher SoPs than PsSo influence which failed to reach statistical significance ($F_{(1,262)}=2.91$ $p=0.09$); the fatigue group showed a highly significant difference with higher PsSo influence ($F_{(1,222)}=9.2$ $p=0.003$); finally the external pain group showed apparently matched PsSo and SoPs effects.

Figure 17 Directional Granger Causality – Physical Symptoms
PsSo = Psycho-somatic causation, SoPs = somato-psychic

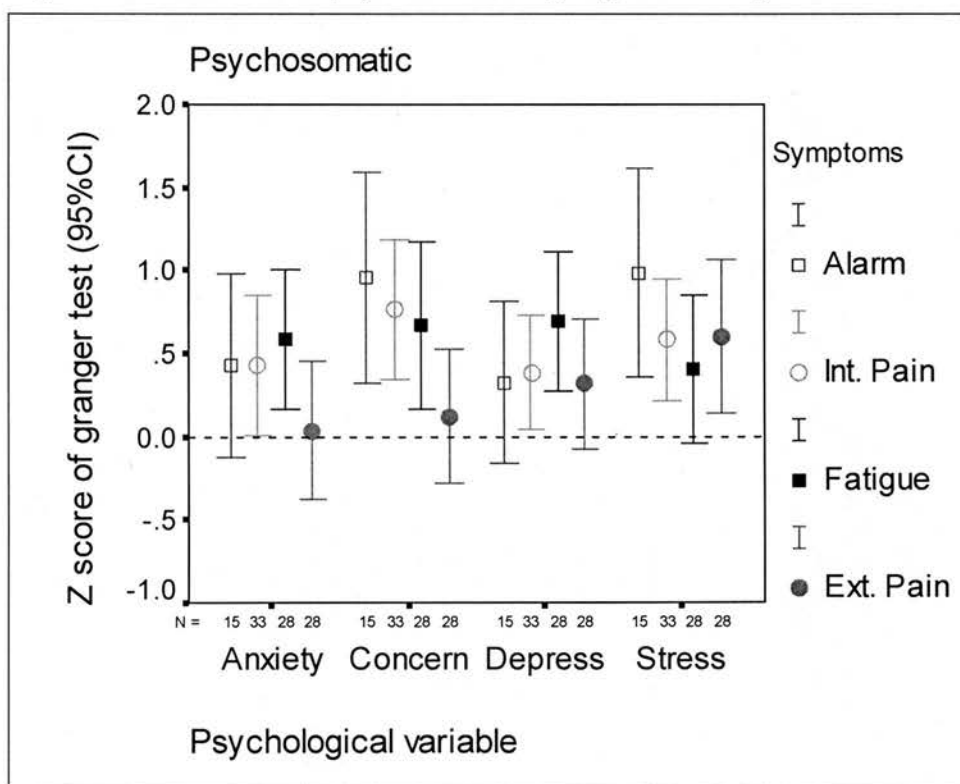


The two pain clusters appeared to behave differently in the granger analysis. External pain had significantly lower Z scores in the PsSo direction than other symptoms

($F_{(1,414)}=6.64, p=.01$) while internal pain z scores were significantly higher than other symptoms in the SoPs direction ($F_{(1,414)}=22.8, p<.001$)

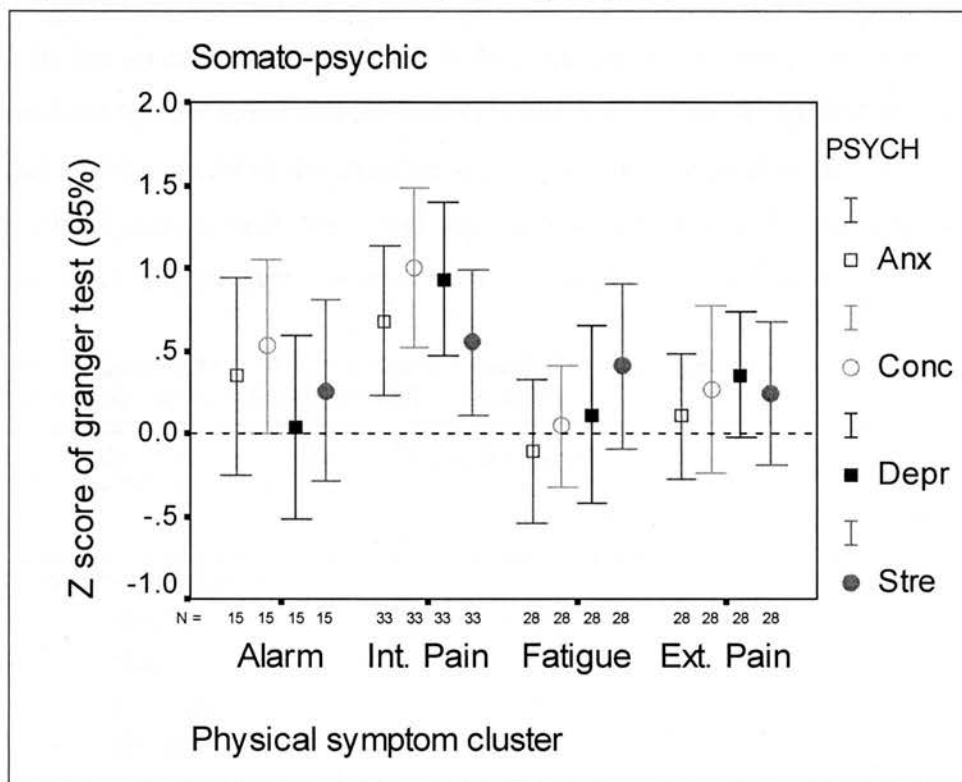
Figure 18 and Figure 19 show more detailed analysis of granger causality results for different psychological and physical symptom pairs. Although confidence intervals are generally wide, a number of trends are apparent. In terms of the PsSo direction (Figure 18), fatigue and internal pain groups appear to be caused approximately equally by all psychological variables while alarm symptoms are directed more by stress and concern than depression or anxiety and external pain seems more influenced by stress than by anxiety or concern.

Figure 18 Breakdown of psycho-somatic granger causality



In the SoPs direction (Figure 19) internal pain appears to have the largest influence on depression and concern, while for all other physical symptom clusters the confidence intervals overlap zero.

Figure 19 Breakdown of somato-psychic granger causality



Influence of baseline characteristics

The granger Z scores were compared between high and low scorers on the baseline measures. There were no significant differences between the baseline groups in mean granger Z scores for anxiety, depression or illness worry. The somatisation group showed an unexpected difference in that high scorers on the SSI showed lower PsSo and SoPs Z scores than the low scorers: this was most marked for fatigue while there was no difference at all for physical alarm symptoms. The difference reached clear statistical significance ($F_{(1,414)}=11.6, p<.001$) in the psychosomatic direction and borderline significance ($p=.049$) in the somato-psychic. Why sequential effects should be less pronounced in those scoring higher on the Whitely 7 scale of somatisation is not clear, however it is unlikely simply to be due to more prominent

concurrent effects drowning out the sequential ones, as the correlation coefficients in Chapter 13 were no different between baseline groups on this measure (in contrast to all the other measures where high baseline tended to be associated with higher correlation).

To examine the distribution of high granger Z scores the highest scoring 10% of results were tabulated and are shown in Table 17. They show a fairly evenly distributed set of interactions in the PsSo direction but the SoPs subgroup is dominated by low mood and the internal pain cluster . The 83 highest granger z scores came from 24 of the 26 subjects, most contributing between 1 and 3, but with two participants contributing 7 and one 13. None of these higher scoring individuals stood out during interview or from their data as particularly unusual.

Table 17 Distribution of interactions yielding high granger scores
cross tabulation of the highest 10% of scores

Category	Physical symptom group				Total
	Alarm	Int. Pain	Fatigue	Ext Pain	
Psychosomatic (PsSo)					
Anxiety	1	3	4	2	10
Concern	4	5	2	1	12
Depress	1	1	3	2	7
Stress	3	4	1	3	11
Total	9	13	10	8	40
Somato-psychic (SoPs)					
Anxiety	1	4	2	1	8
Concern	1	7	0	3	11
Depress	0	9	4	2	15
Stress	0	5	3	1	9
Total	2	25	9	7	43

Conclusion

The granger causality model showed differences between psychological variables and physical symptoms in their interactions with each other. While some findings were in keeping with prior models, for instance the association between current stress and future “alarm” symptoms characterised by altered autonomic function, such as

heart rate and upper gut motility, others were relatively unexpected: particularly the differences between “internal” pain (headache, bowel discomfort, abdominal pain and pelvic pain) and “external” musculoskeletal pain in the extent to which internal pain was subsequently associated with emotional distress and the degree that external pain was relatively uninfluenced by most anxiety, mood and concern, but not stress.

Because of the clustering needed to generate statistical power, direct comparisons with the correlations of differenced data have not been made. It appears however that anxiety and stress have more effect on future symptoms than mood, whereas mood’s main effect is contemporaneous.

This chapter set out to test a null hypothesis: that there would be no detectable causal sequences, however several were identified using an established statistical technique and heterogeneous patterns of mind-body interactions were clearly demonstrated.

Chapter 15 Results - changes related to consultations

Introduction

This section describes the analysis of the changes in symptoms around the time of specific events including consultations. In particular it addresses the hypothesis that *Time series data will show an increase in concern and / or symptoms prior to GP consultation, in keeping with a previously published model*. This was based on work by Cameron & Leventhal who demonstrated a shift in attribution of symptoms from stress to illness prior to consultation(89). In addition the effect of consultations on symptoms was tested to see if these were reduced after the consultation.

Data and methods

For each consultation noted in the electronic diary, three short subsets of data were identified representing baseline, pre-consultation and post consultation phases. Baseline data comprised the seven days from 10 to 4 days before consultation; pre-consultation the 3 days before consultation; and post-consultation the 3 days after the consultation.

98 consultations occurred for which data was available from 2 days before to 3 days after to calculate the effect of consultation. For 80 of these, data was also available from 10 days beforehand; for the remaining 18 consultations, there were missing data which made analysis over the longer period impossible.

The method used was a comparison of within –subject means for the baseline and pre-, and the pre- and post- conditions. Each variable was analysed independently using a paired sample t-test.

Comparison of baseline and pre-consultation

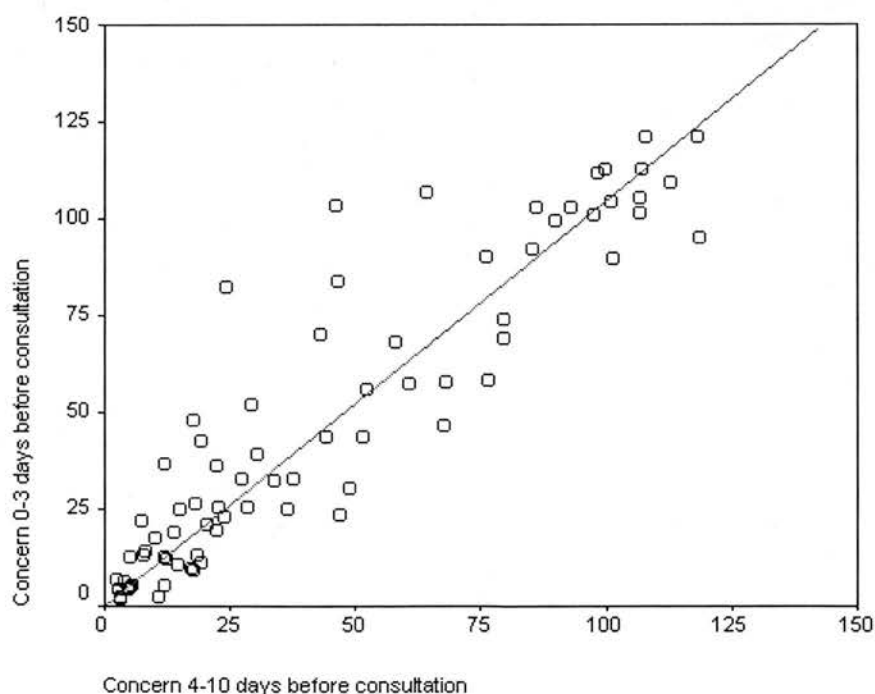
Table 18 shows the results of a paired sample t test comparing baseline and pre-consultation values. There is a significant increase in the mean of concern about symptoms over baseline of 4.5 (95%CI 1.2 - 7.9) points from a rating of 43.4 to 47.9 ($t=2.67, p=.009$). However examination of the scatter plot comparing the two time periods in Figure 20 shows that much of the effect is accounted for by a small number of consultations which yield a very large reduction in concern. These

outliers, while likely to be genuine results rather than measurement error, account for almost all the effect: removal of the five consultations with the greatest fall in concern reduced the difference in means to a non-significant 1.79 ($t=1.4$, $p=.154$). Each of the five of these occurred in different individuals, who had no obvious common characteristic in terms of symptoms or baseline questionnaire measures.

Table 18 paired value t tests comparing pre-consultation with baseline

	Paired Differences			T	Df	Sig. (2-tailed)
	Mean	95% Confidence Interval of the Difference				
		Lower	Upper			
Symptom A	.89	-3.57	5.36	.399	78	.691
Symptom B	.48	-3.49	4.46	.243	78	.809
Symptom C	2.69	-2.44	7.82	1.044	78	.300
Fatigue	1.46	-3.25	6.16	.617	78	.539
Stress	-3.10	-6.44	.24	-1.849	78	.068
Depression	1.23	-2.22	4.69	.711	78	.479
Anxiety	-.24	-4.33	3.84	-.119	78	.905
Concern	4.55	1.16	7.94	2.671	78	.009

Figure 20 Scatter plot of concern in baseline and preconsultation period



Comparison of pre- and post-consultation

The hypothesis that symptoms would reduce after consultation was also tested using similar methods. Comparing means using the t test revealed no significant changes as shown in Table 19; indeed average values for depression and fatigue both rose a little. However as stress and concern had the largest, although non-significant fall (4.17 and 1.7 points respectively) these were plotted to assess the overall pattern as shown in Figure 21 and Figure 22.

Table 19 paired value t tests comparing pre-consultation with post-consultation

	Paired Differences			t	Df	Sig. (2-tailed)
	Mean	95% Confidence Interval of the Difference				
		Lower	Upper			
Symptom A	-3.66	-8.22	.89	-1.60	73	.113
Symptom B	-.76	-5.01	3.48	-.36	73	.722
Symptom C	.09	-5.71	5.90	.03	73	.974
Fatigue	-3.65	-9.54	2.24	-1.23	73	.221
Stress	4.17	-.64	8.99	1.73	73	.088
Depression	-3.98	-9.56	1.60	-1.42	73	.159
Anxiety	.36	-4.40	5.12	.15	73	.879
Concern	1.70	-2.15	5.55	.88	73	.381

Figure 21 Scatter plot of stress in pre-consultation and post-consultation period

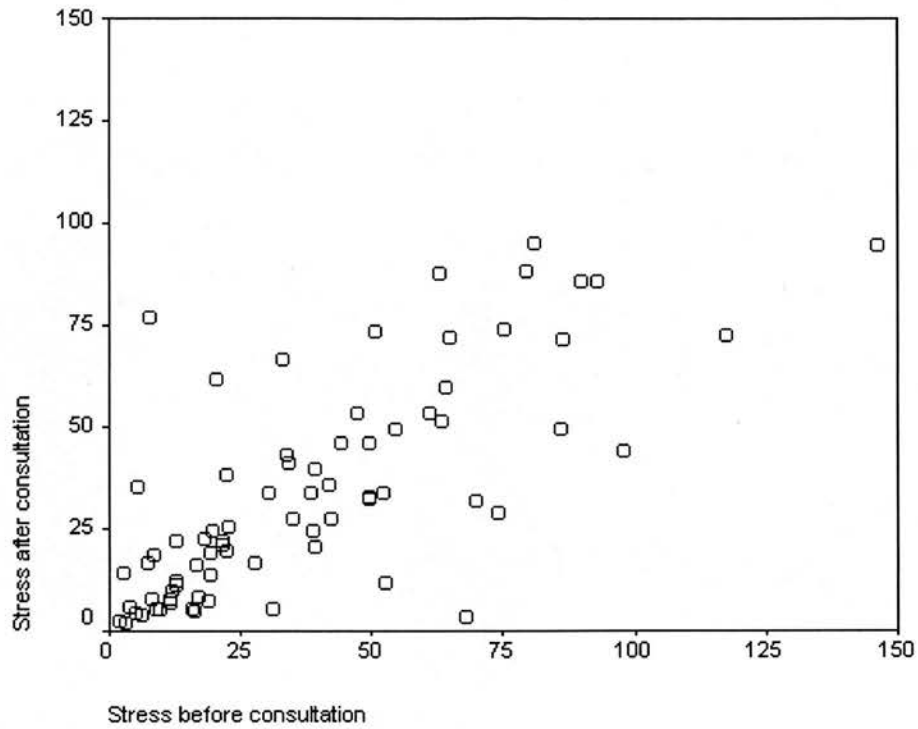
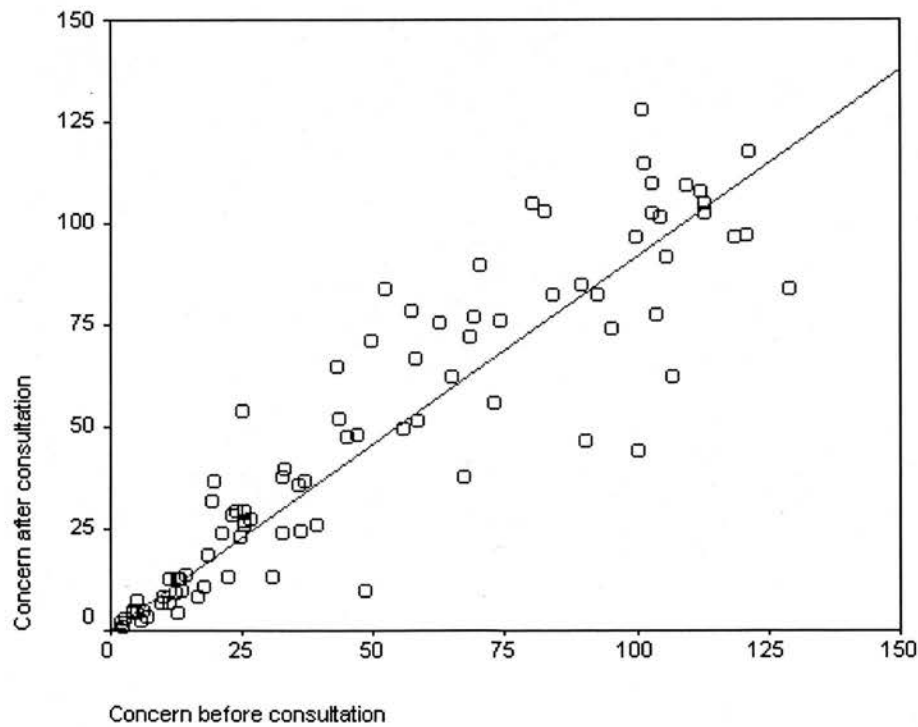


Figure 22 Scatter plot of illness concern in pre-consultation and post-consultation period



Conclusion

While the data show a few instances in which increased concern preceded consultation, these were the exception rather than the rule, and when these were excluded from the analysis there was no detectable effect.

This finding appears to refute the hypothesis that concern would be increased, and contrasts with Cameron's findings(89) from a different setting of new symptoms during a time of life stress. However two factors may obscure this effect in the current research. Firstly there was no way of knowing whether consultations were triggered by distress or by another reason, such as a planned check up or the need to renew a prescription; secondly the participants had long established chronic symptoms and it may be that it would have been possible to observe greater changes before and after consultations with more acute symptoms.

Chapter 16 Results - analysis of complexity

Introduction

This section describes the test of the experimental hypothesis that patients with medically unexplained symptoms have a demonstrably reduced loss of complexity in the day to day variability of their mood and symptoms. This was tested by the calculation of sample entropy (SampEn), a measure of the short term predictability of sequential changes in a time series. The hypothesis proposes that the sample entropy statistic from the data should be lower (indicating greater predictability and therefore less adaptability) than that obtained from a set of corresponding surrogate time series generated using a bootstrapping method.

Sample Entropy Method

For analysis of sample entropy, the paired day time series for each subject and each variable was prepared by setting the time series length to either 64 or 128 points. This was done by either truncating or padding (by adding on data from the beginning of the series) the series to whichever of the two values was nearer the actual length. Paired data were used to maintain diurnal patterns. The length adjustment was necessary for the surrogate generation programme which uses a fast fourier transform of the data into the frequency domain requiring a series length equal to a power of 2. Sample entropy calculations were carried out using a handwritten computer script in the python language. This was checked against the results of an implementation of sample entropy in the R language (Heath RA personal communication). It was designed to run the calculations on batches of files to minimise the risk of operator error in use. The computer script is included in Appendix J.

Generation of surrogate time series and sample entropy distributions

200 surrogate time series were generated for each series of data using the surrogates method of the Tisean software toolkit for non-linear time series analysis(225;226). This generates surrogate data with corresponding mean, variance and trend characteristics, such that any remaining difference between the original time series

and its surrogates is due to intrinsic temporal structure rather than more general properties of the unordered data.

For each of the 200 surrogate series for each subject-variable combination, the SampEn statistic was calculated. These distributions were used as reference populations for their respective data SampEn values.

The surrogate SampEn distributions were found to be approximately normal, and approximately two thirds of a random sample of distributions met strict, Wilks – Shapiro test, criteria for normality.

The actual data SampEn was compared with the surrogate distributions, firstly by comparison with the mean of the respective surrogate distribution and secondly by noting the ranking position of the actual data and allocating a probability value p as this position divided by 200, the number of values in the surrogate distribution.

Where the actual data yielded the lowest value it was accorded a p value of 0.005, the same as if there was one surrogate lower than the actual value. Before analysis, p values were transformed to Z scores using a normal distribution function.

Distributions of sample entropy results

The results of the 208 sample entropy results had a skewed distribution which failed tests for normality (Wilks Shapiro $W= 0.97$, $p<0.001$) with a mean of 1.43, median of 1.54 and range 0.19 -2.83. A similar distribution was plotted for the mean of the 200 surrogate sample entropy calculations generated for each of the 208 series. The group mean of the surrogates was 1.60, median 1.76 and range 0.18 – 2.77. Both distributions are shown in Figure 23.

Figure 23 Histogram of sample entropy scores

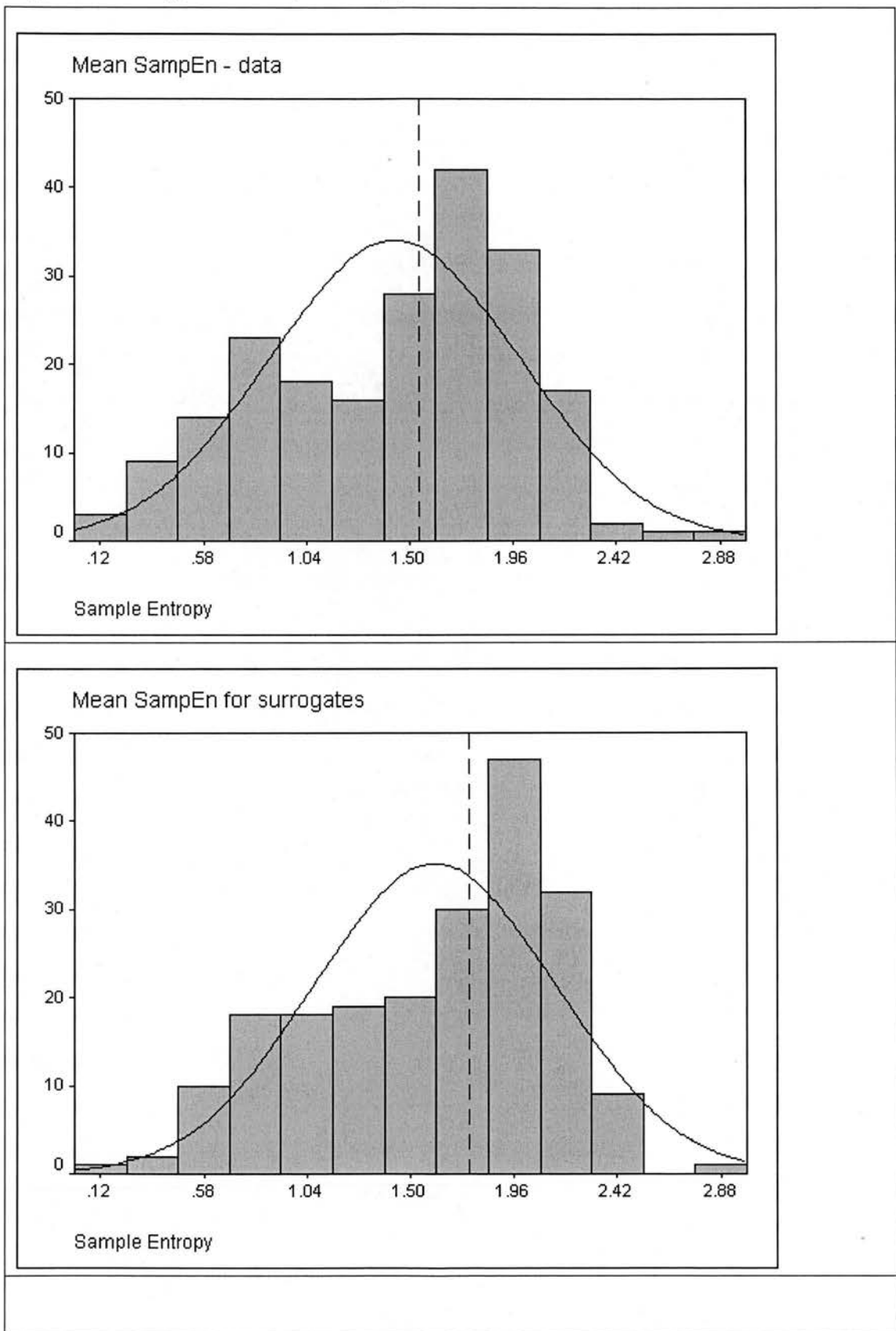
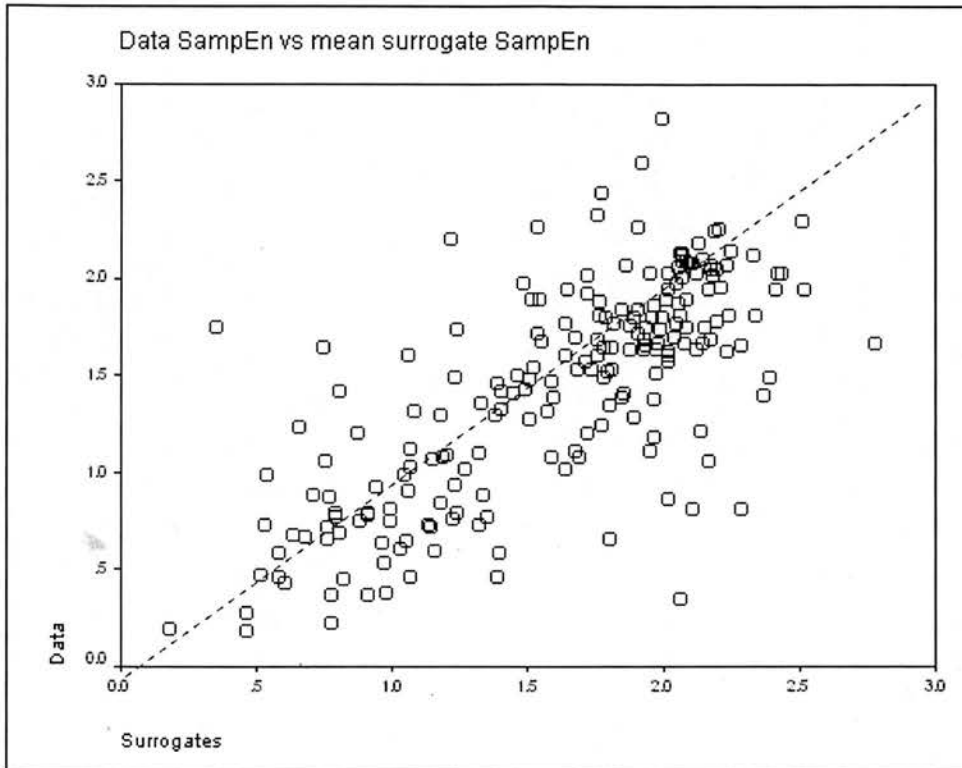


Figure 24 Scatter plot of actual sample entropy of data against mean sample entropy from surrogate data



Comparison of mean actual and surrogate sample entropy Figure 24 shows the scatter plot of actual sample entropy values for data against the mean of the surrogate scores. It shows good correlation but with more scores below the line of equivalence, indicating a lower sample entropy for the actual data compared to the matched mean surrogate value. The correlation appears linear across the range of sample entropy scores.

Comparison of the actual score and surrogate scores for sample entropy was carried out non-parametrically and showed that the two distributions were significantly different (Wilcoxon $Z = -6.679$, $p < .001$).

Subgroup analysis of the mean SampEn for each of the five common variables was carried out using a similar non-parametric comparison against pooled surrogate means. The results of the Wilcoxon signed rank test are shown in Table 20, demonstrating that entropy was significantly lower in all symptoms than for surrogate data

Table 20 Mean Sample Entropy for data and surrogates by common symptoms

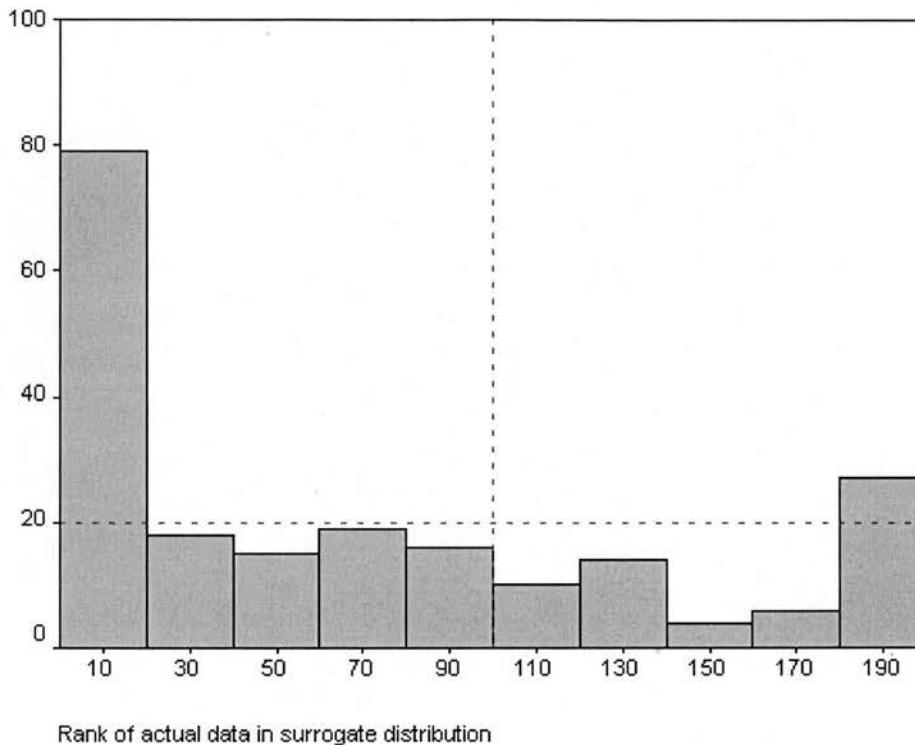
Mean Sample Entropy value				
Variable	Data	Surrogates	Wilcoxon Z	p
Anxiety	1.56	1.74	-2.9	.004
Fatigue	1.24	1.40	-2.43	.015
Depression	1.59	1.75	-2.9	.004
Concern	1.34	1.53	-2.58	.010
Stress	1.30	1.48	-2.48	.013
All	1.44	1.61	-6.68	<0.001

Ranking of actual sample entropy within surrogate distributions

Figure 25 demonstrates the distribution of the rank within the respective surrogate distributions of the actual SampEn values. It clearly demonstrates a clustering at the low end of the distribution with 79/208 (38%) of values ranked in the lowest 10% of the distribution. The horizontal line on the graph demonstrates the expected number in each bin of the histogram if the actual SampEn values were randomly distributed through the surrogate distributions. The graph also shows a small excess of values at the upper end of the distribution.

Figure 25 Histogram of ranking of actual sample entropy within surrogate distributions.

Horizontal dashed line indicates level of even spacing, vertical dashed line indicates median value



Conversion of sample entropy & ranks to Z scores

In order to facilitate further comparisons, probability (p) values, derived by dividing the ranking score by the total number of surrogates (200), were converted to z scores using a normal distribution function, whereby a z score of +1 and -1 represent one standard deviation above and below the mean of a normally distributed population.

Figure 26 shows the distribution of z scores derived from the ranking against surrogates. There is considerable clustering at both ends of the distribution owing to the restriction imposed by calculating 200 surrogates (limiting the Z scores to the range -2.75 to +2.75). This end effect can be transformed using calculated z scores derived from the mean and standard deviation of the each set of surrogate data by assuming that the each surrogate set is normally distributed. The results of this, showing a smooth curve but very distant outliers is shown in Figure 27. The vertical dashed lines in Figure 27 represent the total range shown in Figure 26.

Despite the obvious end of range clustering, the ranking derived set of z scores was chosen for further comparison as it was less dependent on assumptions of normality in the data. Further analyses involving the z scores were carried out using non-parametric methods.

Figure 26 Z scores for sample entropy

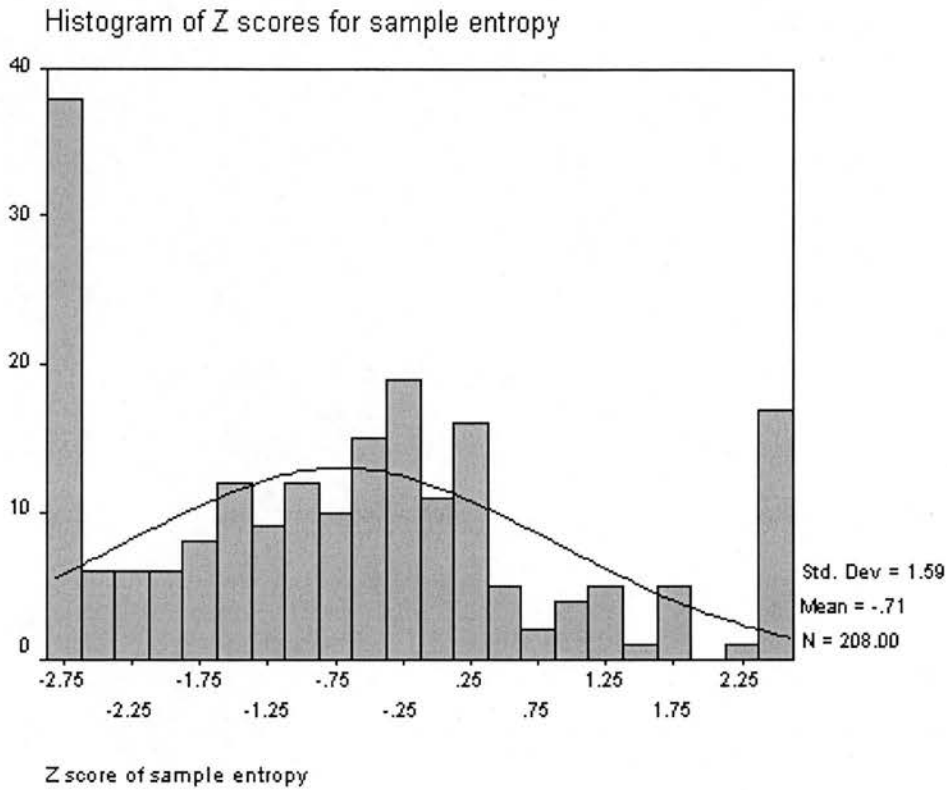
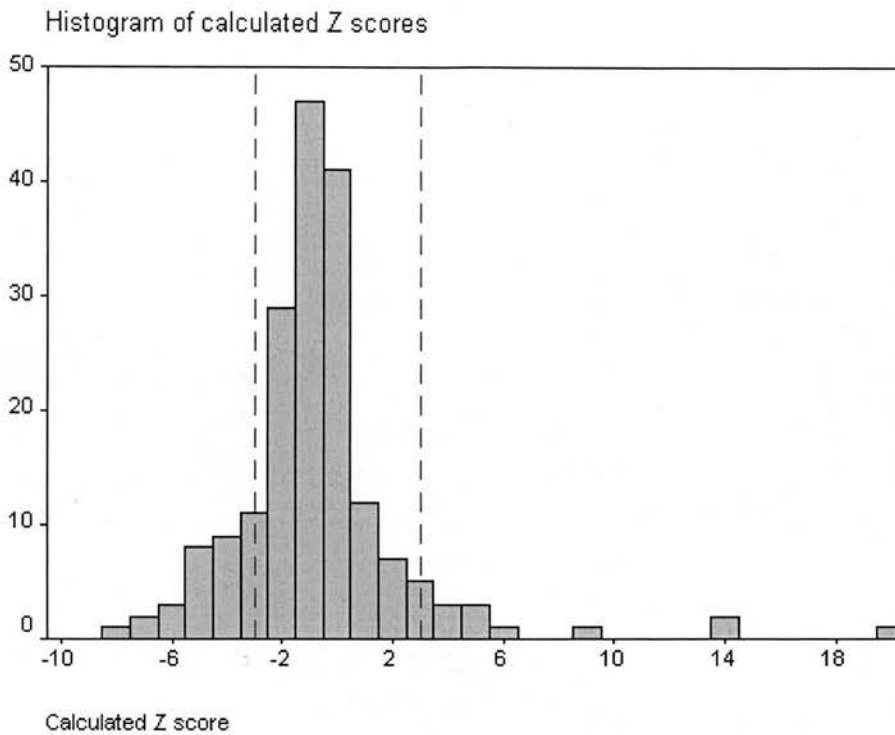


Figure 27 Histogram of Z scores obtained by calculation



Z scores

The mean z score was -0.71 (N= 208, SEM 0.11). Z scores and estimated 95% confidence intervals were calculated for each of the psychological symptoms and for fatigue and were significant for all except stress (see Table 21).

Lower entropy (i.e. more negative Z score) is indicative of greater repetition of patterns of short sequences and is assumed to mean less reaction or adaptation to changes in the environment. In this instance one would expect stress (which should be a state function of the environment) to have a higher entropy than anxiety, if that was more of a trait variable and this is the case.

Table 21 Z score for sample entropy, for each of the five symptoms recorded by each participant

Variable	Z score	(95% CI)
Anxiety	-1.01	(-0.4 to -1.6)
Fatigue	-0.77	(-0.22 to -1.31)
Depression	-0.65	(-0.14 to -1.16)
Concern	-0.80	(-0.15 to -1.46)
Stress	-0.53	(0.11 to -1.19)
All	-0.75	(-0.49 to -0.91)

The mean Z scores for the participants' chosen physical symptoms were less negative (i.e. relatively higher entropy) than for either fatigue or for the pooled psychological variables and are listed in Table 22.

Table 22 Average sample entropy z score for physical symptoms

Symptom	Number of cases	Average Z score of sample entropy	
Musculoskeletal			
Muscle pain	15	-0.59	
Joint pain	12	-0.67	
Back pain	1	2.57	
	28		-0.51
Digestive			
Bowel Symptoms	10	-0.48	
Abdo pain	9	-0.59	
Dyspepsia	5	-0.29	
Nausea	3	-1.44	
	27		-0.59
Headache			
	12	-0.44	-0.44
Other			
Throat tight	3	-1.84	
Chest pain	2	-2.57	
Pelvic Pain	2	-0.56	
Weakness	2	-2.16	
Numbness	1	2.58	
Palpitations	1	-0.60	
	11		-1.28
Total	78	-0.64	

Sensitivity of sample entropy to large scale properties of the data

Because sample entropy depends on comparing short subsequences of a time series from different positions in the data it is sensitive to large scale changes such as trends in mean or variance over the time of the series. The method of generating surrogates should also generate these trends such that the z scored SampEn statistic should not be influenced by trends or absolute measures of either mean or variation.

To test for this, the same method for estimating trends and absolute values for the mean and standard deviation was used as in the earlier section on properties of the time series, in which trend was taken as the difference in both mean and SD between the first 40 and the last 40 readings of each series and the value of both over the last 40 readings was used as the final mean and SD. Correlation coefficients were calculated between sample entropy and the value and trend for the mean and SD of each data series and are reported in table All correlations were significant (r between 0.15 and 0.48) suggesting that the sample entropy score was dependent on large scale characteristics of the data.

The use of z scores effectively removed the correlation with values and trends for mean and SD of each series, however there were a small number of outliers with very low sample entropy z scores and large trends as shown in the scatter plot figure. Table 23 summarises the results of these correlations.

Table 23 Correlation coefficients between sample entropy or z score and large scale characteristics of time series

	Mean	Trend in mean	Standard Deviation (SD)	Trend in SD
Sample Entropy	.483(**)	.255(**)	.151(*)	.191(**)
Z score of Sample Entropy	.072	.099	.061	-.033

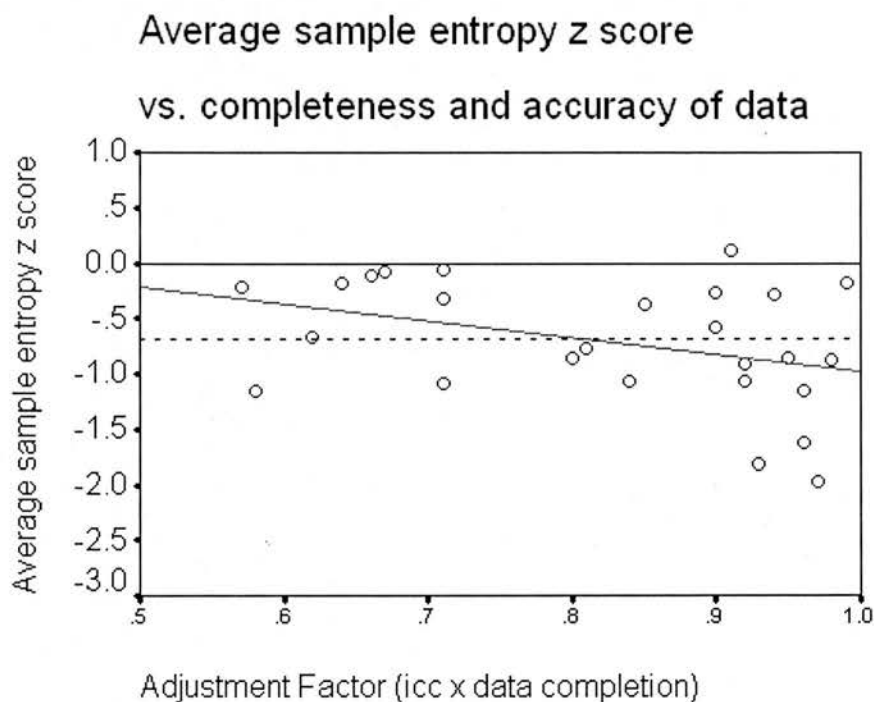
** Correlation is significant at the 0.01 level (2-tailed).
 * Correlation is significant at the 0.05 level (2-tailed).

As using z scores for SampEn appeared to remove the statistic's sensitivity to large scale characteristics of the data, all further entropy analysis was carried out using the Z score for each series.

Sensitivity of Sample Entropy to data quality

As SampEn depends on the identification of patterns in the fine grained detail of the data, it may be susceptible to error in data collection. Figure 28 demonstrates that the trend is for SampEn to be further away from the average in data from subjects with a higher data quality, here represented by the same adjustment factor as used in the meta-analysis of correlations. This supports the argument that the low entropy values seen are not simply a function of poor recording, instead they occur more often in the more accurately recorded data.

Figure 28 Scatter plot of sample entropy against data quality



Sample entropy differences between variables

In order to explore differences between variables in sample entropy, non-parametric (Spearman's rank) correlation coefficients were estimated between z score sample entropy for each of the four psychological symptoms and fatigue. Three moderately strong correlations were seen: between fatigue and anxiety ($\rho = -0.44, p = .025$); fatigue and concern ($\rho = -0.48, p = .012$) and stress and depression ($\rho = -0.43, p = .036$). All three were negative (i.e. high entropy in one was correlated with low entropy in the other). None of these interactions showed more than weak correlations

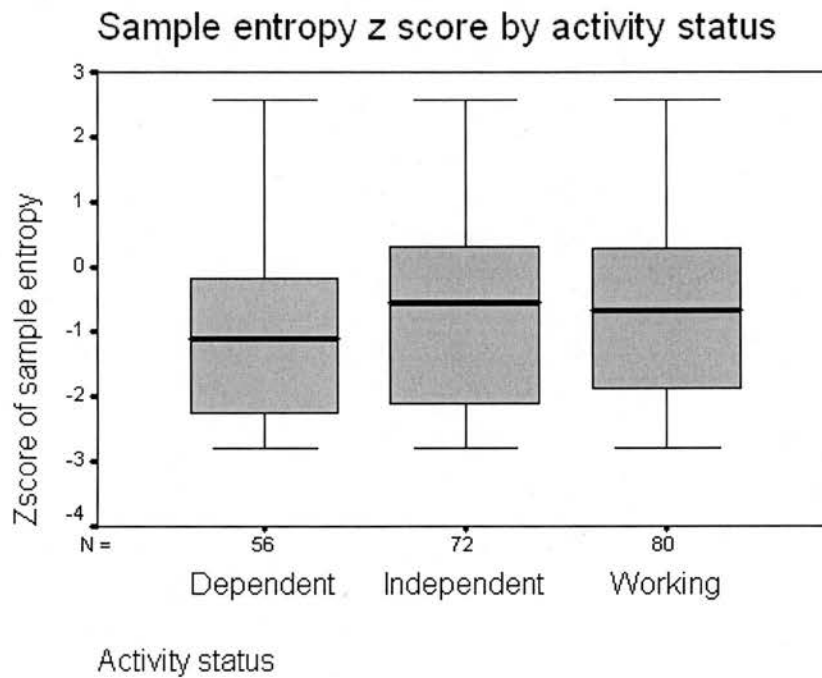
in the conventional analysis of correlations between variables, while none of the strong positive correlations interactions seen in the meta-analysis of simple correlations - such as fatigue with depression or anxiety with stress were significant in correlation of entropy scores.

Comparison of Sample Entropy with baseline variables

There were no significant correlations between the z scores for sample entropy on any variable and the results of the baseline questionnaires.

There was a small difference, which did not reach statistical significance in the distribution of z scores between the three subjective categories of activity level: working, independent but not working, and dependent shown in Figure 29.

Figure 29 Sample Entropy by activity status



Clustering by sample entropy results

An exploratory cluster analysis by SampEn results for psychological variables and mood was performed using the 2 and 3 cluster methods in SPSS. The two cluster model was chosen for further investigation as the two clusters were approximately evenly balanced and represented a difference in sample entropy scores as demonstrated in Table 24.

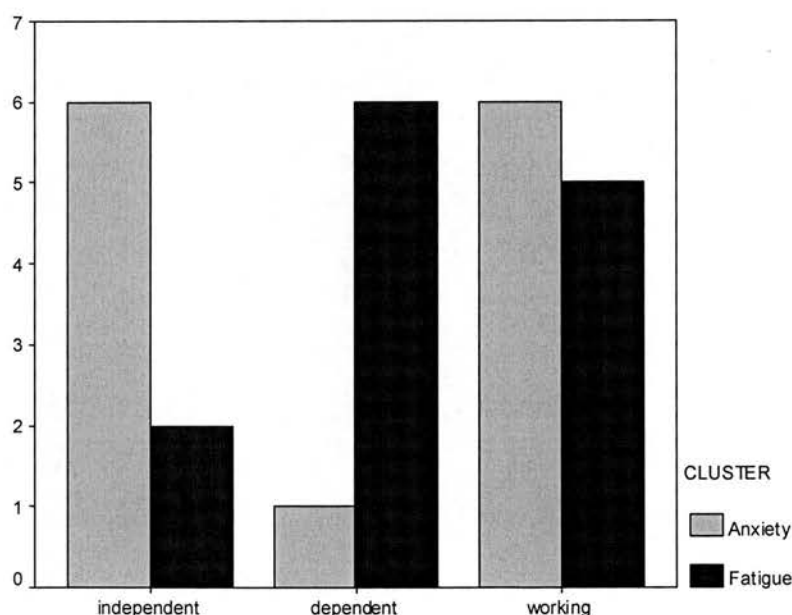
Table 24 Comparison of z scores for common variables in the

Cluster	Z score		F(1,24)	p
	1	2		
Fatigue	0.158	-1.69	23.1	0.000
Stress	-0.785	-0.292	0.60	0.446
Depression	-1.30	-0.959	1.58	0.221
Anxiety	-1.94	-0.079	15.8	0.001
Concern	-1.81	0.210	16.2	0.967

The two clusters were characterised as low anxiety entropy (group 1) and low fatigue entropy (group 2). Despite this, there was no difference between the two groups in baseline levels of anxiety, depression or concern.

There was however a possible difference in activity status of the members of the two cluster groups as shown in Figure 30. This demonstrates that while the two groups were equally represented in the participants currently in employment or seeking work, there was a split between the groups in terms of being independent but not working, or dependent on others for daily activities. The dependent activity group contained 6 members of the low fatigue entropy cluster and one of the low anxiety entropy cluster while the non-working but independent group contained 2 and 6 respectively. This was a post hoc exploratory analysis with small numbers so should be treated with caution.

Figure 30 Activity Status by Sample Entropy cluster



Conclusions

The sample entropy analysis of data and surrogates appeared technically satisfactory. The results showed the actual statistical entropy of the separate time series variables to be significantly lower than the mean of the surrogate data.

By using z scores, a measure of the statistic relative to a bootstrapped set of surrogates, dependence on large scale characteristics of the data such as trend and variance was effectively removed. Anxiety and fatigue had the lowest entropy and cluster analysis suggested two groups of participants characterised by low entropy of one or the other of these.

The finding of reduced entropy, or complexity, is in keeping with the hypothesis that patients with MUS should show this, in association with diminished adaptability and fitness

Chapter 17 Results - interview data

Introduction

This section describes and discusses the content of both the recruitment and feedback interviews with participants in the study. It does so with reference to contemporaneous field notes which were compiled in an unstructured fashion during the interviews by the author. The interviews were not recorded in any other way.

Interviews lasted 10-20 minutes at the first visit so a relatively brief clinical history could be taken, with the remainder of the visit involved in explanation of the study, consent and the administration of baseline measures. The second interviews lasted between 10 and 60 minutes during which the results of the diary and its statistical analysis were fed back to the subject and interpretations of the diary data and patient narrative were explored. This exploration was conducted in a semi-structured way beginning with the participant's perspective on the results, through elaboration of the timeline of the illness and discussions around possible causal factors and triggers. The interviews took the form of a medical consultation in that interpretations were offered and suggestions for action were made as would happen in an extended clinical consultation rather than a pure research interview.

Given that these were consultation – interviews and that analysis was based on limited field notes and recall prompted by them, detailed formal qualitative analysis was not appropriate. The interviews were primarily designed to feed back hypotheses from the data analysis to participants in an evolutionary way to explore the possibility of using electronic diaries in a clinical context and to offer some value to the participants, although it was repeatedly made clear that this was not “therapy”. Nevertheless the field note data were used to examine a number of themes previously identified from the literature to confirm their applicability to this study population.

Usefulness of the diary

Participants were broadly positive about the diary. Most did not elaborate greatly, not least because the first question in an interview which was not clearly mapped out in advance, concerned how they had fared with the diary, so responses were guardedly positive. Nonetheless several commented that it had helped them to disentangle

different elements of their symptoms, for instance recognising that anxiety changed from day to day, or that there were rather more pain free days than originally thought. One participant reported her teenage daughter observing her and disagreeing with one value (mood): the participant came to realise that her family were more aware of her mood changes than she was herself.

Relatively few participants noted specific associations from the diary results and correlations as presented. One drew attention to a sudden spell of more intense pain levels which she clearly recalled experiencing and catastrophising: "I thought I had cancer from the neck down that week". Another saw the link between mood and self-reported gastro-intestinal symptoms and commented that his grandmother had always said she felt her nerves in her stomach.

During three of the post-diary interviews it became apparent that emotional trauma had happened to the interviewee in the past. In each case sensitive and non-directive steps were taken towards the area of concern without pressure. The approach was that which would have been used in a long primary care consultation. Two of the three had spoken briefly to a counsellor previously about the issues but had not followed that up. All were encouraged to seek additional support, in one case the patient's GP was notified, with her consent, the same day by telephone.

Choice of themes

Five themes from the literature were selected for examination: the sufferer as a strong person; the illness story as a chaotic narrative; problems of illness labelling; all or nothing behaviour and the linking of mind and body in explanatory models. The themes were chosen because of a combination of frequent occurrence in the participants' stories (particularly "the strong person" and the search for labels) and from review of recent publications, for instance on patients with chronic pain(27;229;230) and unexplained symptoms(33;231).

These themes, already evident in the literature, were reckoned to be important areas for use in consultations with other patients with symptoms. This was either because they represent the ways patients describe the paradoxes of their own experience (the chaotic narrative, the strong person), because they are a frequent source of disagreement (labels, mind-body explanations) or because they are areas where a

common ground between doctor and patient may be established in order to build a therapeutic relationship (explanatory models, all-or- nothing behaviour).

The strong individual

A number of accounts of patients with functional symptom disorders have drawn attention to the assertion of current or prior strength of sufferers(27). This appears to be partly in defence of the perception that they may be unfairly occupying a sick role in society, but may also have a protective effect on mental health, portraying a strong self in a weak body rather than a weakened self.

Several features of participants' accounts of themselves strongly support this notion of projected strength.

Strength in relation to other illness

Several participants reported that while they were limited by their predominant unexplained illness, they were able to deal perfectly normally with other ailments:

“[My headaches are] not a normal headache, I can't just take paracetamol and carry on” (Subject 44, female, mixed symptoms)

“[Tension type headaches] ... like ordinary people”

“It's easier to live with normal fatigue and joint pains – they pass” (46, female, GI symptoms)

This extended to severe pain

“I always had a high pain threshold.... I walked into a gynaecology ward with an ectopic pregnancy” (65, female, pain)

and an episode of tachycardia requiring hospital admission

“[When the tachycardia is there] I can cope, take things in my stride, just to get on with it”(42, female, fatigue).

These accounts highlight the difference in severity between the current symptoms and “normal” illness.

Strength as an individual

There were a number of indicators that participants saw themselves as strong in spite of their symptoms. For instance one reported that if in pain he could “get on if it's critical” but would suffer the next day.(40, male, pain)

Others described the difficulties they faced each day such as [when experiencing fatigue in the mornings] “I could cry because of the full day ahead.”(42) while others reported tackling symptoms head-on “If I feel bad, the best thing to do is physical activity – then I feel fine except physically tired”(37, male, fatigue).

Several studies, particularly of chronic fatigue syndrome have drawn attention to the idea of high levels of achievement in the premorbid state. Two participants had taken part in competitive athletics before they became ill and several reported trying to go to the gym in the early stages of their current illness.

Strength relative to others

Participants used several different techniques to project their strength in relation to others. Several talked dismissively of other patients with unexplained illness, or of other family members they saw as less strong:

“not like those ME people who just give up”(57, female, fatigue)

“I just get on with it, my sister cries at things”(54, female, GI)

Relatively few called on others’ authority to support them, although one used her husband’s account of her coping strategies.

The chaotic narrative

In analysing the narratives of a group of patients with unexplained neurological symptoms, Nettleton and colleagues(33)identified a characteristic chaotic narrative style which had indeterminate beginning and end and an unpredictable trajectory. It was portrayed as contrasting with the more conventional narrative styles of restitution (in which a problem is identified and then resolved) and the quest, in which the sufferer gains insight and self knowledge through the experience of illness.

Instead patients with MUS were identified in this work as having chaotic narratives which included “embodied uncertainty”(231) and “ambivalence” leading to a perennial questioning of what is wrong with the body and what it means for medicine to be unable to find it.

Unpredictability

Several subjects reported finding their day to day variation in symptoms hard to deal with:

“[It’s the] unpredictability of the bodily aches and pains, ..[they can occur] literally anywhere”(37, male, fatigue)

“I don’t know what to expect, I would like to know..[I don’t like] the uncertainty rather than being prepared”(35, female, pain)

For some, while that is the case now, it was not always so:

“The pain is variable, it used to relate to doing things, but not now. Sometimes I wake up in the mornings and it’s there”(55, female, fatigue)

But a degree of acceptance of unpredictability was also possible:

“Initially I didn’t accept what was wrong, now I accept that I can’t influence it, but I know what it is”(65, female, pain)

Making sense of unpredictability

Participants appeared to respond to unpredictability in different ways. For some it seemed to be interpreted as malign fate, or expressed as a simple wish for order (quickly countered by the defence that such a wish did not mean obsessionality).

Others sought to find explanations:

“There’s always got to be an explanation, yes?”(36, female, mixed symptoms)

Others tried to find links, such as environmental triggers or digestion, however it was clear that such explanations broke down:

“The back pain would be connected to the neck... but they can’t all be connected”(37, male, fatigue)

The burden of unpredictability

The alternative to making sense of the illness or passivity in the face of it, appeared to be to continue life with an ongoing vulnerability:

“ No control at all... out of control and it’s going to happen again”(45, female, GI)

“[You don’t] make sense of it, it’s there, you just have to go on. Smile on the outside. The alternative is to give up”(54, female, GI)

For these people, the chaotic narrative meant an additional burden, on top of the experience of symptoms. One subject recognised that other people too were impotent in the face of the illness, expressing sympathy for her GP who “tries but has nothing else he can do”, returning for appointments even though her husband thought it a waste of time and chided her for doing so.

The work of making sense

A number of participants described their search for an explanation in more positive terms as a task on which they were engaged.

“I don’t like unexplained things, I have intelligence and should seek out the answer; I don’t like to be beaten”(37, male, fatigue)

The attempt to understand is “fascinating...it becomes your whole life”(49, female, fatigue)

This quest was portrayed in very different terms from the caricature of neurotic self-obsession used by doctors of their “heartsink patients” as both a moral duty and a worthy intellectual challenge. It may, however, be that this approach combines both the outward projection of participation in a worthwhile quest with a distraction from facing deeper issues.

Other narrative styles

Despite the unexplained nature of symptoms, other participants appeared to have transformed, or be transforming their narratives into other styles. The work of making sense can be seen as a strand of the quest narrative. Others more clearly described a restitution narrative: of the journey of recovery, more or less complete, from turbulence and unpredictability in their life towards more settled times.

Ambivalence over labels

While having a label or name for a condition or symptom is seen as important in belief models(232;233) , Nettleton noted an ambivalence about illness labels in which there was a tension between having and not having a label for one’s illness(231). Recent work on fibromyalgia confirms this ambivalence extends to an initial relief that a label exists being replaced by a return to the struggle to make sense of the diagnosis(234). Comments recorded in the field notes demonstrated this along with related issues of authority over labelling.

Some participants clearly expressed a wish for resolution of the uncertainty, which was not their fault:

“They’re at a loss, I’m looking for a name; doctors never said what to do”(51, female, GI)

“ [The failure of doctors] to find something depresses me; I wish they’d find something, anything”(54, female, GI)

Here the implied theme is that anything, even perhaps a serious illness, would be better than uncertainty

Ambivalence over labels for physical diagnoses

While patients with MUS are often stereotyped as anxious for an explanation several participants were clearly weighing their own symptoms up against a medical label:

“[I get] alternating constipation and diarrhoea: I don’t know if it’s IBS” (61, female, fatigue)

(describing idiosyncratic swings in the amount of daily energy) “that’s why I know it’s not ME”(49, female, fatigue)

Often this ambivalence was expressed in negative terms of what the label can’t be:

(describing neck and shoulder pain) “I’ve never had whiplash”

Ambivalence about the label was strongest for ME (Myalgic Encephalomyelitis)

One reported wondering “Am I giving myself ME?”(46, female, GI)

Another screwed up her face as she introduced the label of ME as the one that best fitted her condition(56, female, pain). And a third:

“[I was told] “you’ve got ME”. ME is what they use when they don’t know what’s wrong with you, it’s not a real disease” (57, female, fatigue)

Despite this initial resistance this participant maintained the label as it gave her access to other forms of support and benefits which came with it.

Ambivalence was clearly not limited just to patients. One reported:

“My GP thought it was depression, but then reluctantly suggested fibromyalgia, ‘but that’s usually crackpots’ she said”(58,female, pain).

Authority and defiance

Some used themes around labelling to express authority in deciding illness labels, either locating it with the doctor:

“[I] call on the authority of what the doctors have said “(36, female, mixed)

or in defiance of what was seen as medical mis-labelling:

“My GP said my stomach pain was not an ulcer, but the endoscopy showed small ulcers even though the hospital doctors said it was stress”(58, female, pain)

One in particular sought out a specialist to confirm her personal (correct) diagnosis of fibromyalgia and another reported the GP who gave her a diagnosis of ME as “wonderful”.

In contrast with the ambivalence or negative issues around labelling, a few participants reported much more positive experiences, largely in relation to diagnosis of fibromyalgia in association with a community support group for sufferers from the condition. Here personal ambivalence could be cast aside for group security and shared ownership of difficulties and distress in keeping with Madden’s argument that the value of the fibromyalgia label depends on the context within which it can be used to create meaning(234).

Absence of ambivalence over psychiatric labels

In contrast with ambivalence over physical illness labels, psychiatric labels for symptoms were seen as more clear cut. Several participants were concerned that people saw them as hypochondriacs or mad. None believed they were, and this concern over the social projection, rather than internal doubt, seemed to be the key here.

Labels of depression were often refuted:

“I didn’t have depression, my mother had depression she felt nothing. She wasn’t bothered and was going to kill herself”(57, female, fatigue)

“I’m not depressed, I’m frustrated and I get angry”(64, female, pain)

or were explained as secondary to the uncertainty or the symptoms themselves:

“I’m depressed, because they can’t find anything” (45, female, GI)

“pain drives the depression”.(42,female, fatigue)

When acknowledging psychiatric labels, this tended to be in the past, before their current illness:

“Lots of people said it was stress, but I had stress before and had got over it [before the symptoms started]”(60, female, GI)

“I had postnatal depression after the birth of my second child, but I’m not sad now”(61, female,GI)

All or nothing behaviour

The concept of “all or nothing behaviour” describes a tendency for the individual to keep going in the face of difficulty such as illness(235). This is not just subjective, but can be objectively measured, for instance in terms of time taken off work.

In a recent prospective study of the onset of irritable bowel syndrome following proven campylobacter enteritis, the all or nothing personality component of the Behavioural Response to Illness Questionnaire, independently predicted the perseverance of symptoms, in contrast to the other coping styles (limiting behaviour, practical help seeking and emotional help seeking)(235).

All or nothing at the onset of illness

Several participants reported trying extra hard to overcome their symptoms, particularly in the early stages: one for instance persisting in walking up hills, even going sideways when myalgia became severe, and going to the gym to try and get fit despite pain. Another noted that

“the more I did to get fit, the more the headaches and pain got worse”(56, female, pain)

while one reported that he

“kept going when the symptoms started, despite increasing fatigue and abdominal problems”(33,male, fatigue)

and another was

“working at the outset, feeling exhausted, still going to aerobics” (61, female, fatigue).

Adapting to all or nothing

Towards the end of the study, as all or nothing type behaviour had been reported by several participants, some interviewees were given the phrase to see how it applied to them. For instance one reported how she used to

“tear along, setting and meeting deadlines... I liked to be in control...I wasn't good at delegating... [and having left that environment because of her health] ..I came here to escape that. (Subject 65, female, pain)

However, another participant used the thinking behind all or nothing behaviour against herself

“sometimes doing more gets rid of it (fatigue), therefore [I think] there's nothing wrong” (Subject 55, female, fatigue)

All or nothing in relationships

Several participants also seemed to display all or nothing type behaviour in their interpersonal relationships, and to take comments by others including professionals very literally. One, who had multiple dysfunctional relationships and marked physical disability also reported being told in her 20's "if you keep lifting like that you'll harm your back and be crippled when you're 50"(32, female, pain). She appeared to regard all that had happened since as the fulfilment of this prophecy.

One can speculate that this lack of subtlety may overlap with the extensively measured concept of alexithymia(236), which relates to difficulty in communicating emotions to others. While this was not formally tested in the study, a recurring impression was that this may well overlap with all or nothing thinking and this may be worth further investigation.

When the term "all or nothing behaviour" was offered to participants it was consistently well received. It was seen as an attribute which they could recognise in themselves which was both morally good and also provided a justifiable vulnerability to illness.

Linking mind and body in explanatory models

The interviews demonstrated that most participants were not averse to linking mind and body in their symptoms. However such explanations were seen as difficult, for fear that they may be seen as implying a psychological cause.

The concern that it might be all in the mind

Several participants expressed the view that they feared that others might think their illness was "all in the mind" or that they were imagining it but made it clear that this was not the case. A smaller number explicated their ways of knowing this, describing personal experiments which showed that physical things worsened their physical pain, therefore it was not psychological.

Almost all description of "all in the mind" attribution was in the social domain, that is it was phrased in terms of how others might perceive this. Only a couple of the participants described psychological vulnerability independent of their symptoms including one who, by the second interview appeared fairly clear that her symptoms were secondary to an anxiety state which was in turn secondary to family stresses.

Refutation of causality

The problem of causality recurred throughout the interviews. While there is much published work on the types of attributions people make for illness(34;51) there is little on the analytic tools and methods they use in making them.

In particular, interviewees seemed to use a refutational logic based on recall of feelings and the notion of necessary cause. Hence in order for stress to cause an illness they must have felt stressed at the time the illness began. Thus:

“I wasn’t depressed when I got it (the illness), I have been since, the constant tiredness then demoralisation” (57, female, fatigue)

ruled out depression, or even a potentially depressogenic environment, at the time of onset. This refutational logic appeared more important in dealing with the fundamental nature of the name of the illness, whereas greater flexibility was allowed with day to day triggers where an association rather than a necessary cause appeared sufficient. However this was not always the case:

(when asked about headaches and tension) “I don’t believe that, I can be calm”

The difficulty of finding acceptable interpretations

Several interviewees reported the work(27) of interpreting their illness as being difficult and required the weighing of different forms of evidence:

“I’ve spent a long time trying to work out which of physical symptoms and depression came first”(33, male, fatigue).

“It (correlation between abdominal pain and concern) could be anticipation of a restless night. But I’ve read somewhere that [worry can cause pain]”(48, male, pain)

Simple observations, such as that IBS symptoms could be worsened by worrying, could be helpful; but some interviewees sought models with greater detail:

“Relaxation helped, but there’s no reason why that worked.... [compared with] running works because of the endorphins.”(37, male, fatigue)

Finally some interviewees found interpretations that were helpful, for instance an anxious tightening of the throat was a sign that it was time to slow down.

Conclusion

This section has not explored new topics, rather it has tested the findings of formal qualitative studies against this study population over the two interviews before and after the diary part of the study.

A focus of this confirmatory analysis was on areas of paradox and tension in individuals' accounts of themselves: the contrast of apparent weakness but hidden strength; the ambivalence over disease labels which allowed constraining of their experience by a medical system claiming both to know it but to be unable to explain it; and possible ways of sharing understanding about mind-body linkages. These paradoxes of daily experience may take place in the patients' own lifeworld(237), and it seems likely that for some at least, the ability to deal in the lifeworld, may be both important and valuable.

In the case of the three participants who disclosed past severe traumas, this ability to accept the individual's account and context of symptoms, within a personal narrative and ways of coping, appeared to facilitate trust and a willingness to reveal deeper concerns that could not, previously, be fitted into the larger picture.

Chapter 18 Discussion – review of main findings

This chapter summarises the results as they address the questions and hypotheses outlined in Chapter 6.

Specific Questions and hypotheses

Research Question 1 : Are electronic diaries a suitable tool for symptoms research?

Participants were able to record self-ratings consistently and with good to excellent reliability. Data handling procedures were uncomplicated and little data was lost.

Electronic diaries appear well suited to longitudinal symptoms research studies.

(Chapter 11)

Question 2: How do symptoms and emotional states reported by electronic diary vary over time in patients with persistent MUS?

Hypothesis

Symptoms and emotional variables will vary over time, but not consistently so, either within or between individuals. This hypothesis was confirmed, and in this group of chronically symptomatic individuals, there was little long term trend in the majority of participants.

Symptom ratings will show autocorrelation which can be statistically removed to permit further analysis. Autocorrelation was present in the data as hypothesised and could be statistically dealt with. Beyond autocorrelation, there was no consistent pattern which could be easily visualised without statistical analysis. (Chapter 12)

Question 3: What are the concurrent associations between symptoms and mood?

Hypothesis

There will be significant associations between physical and emotional variables.

Associations will differ within and between individuals. This was demonstrated to be the case, particularly for mood, which was moderately strongly correlated to several concurrent physical symptoms, and for concern which was correlated with visceral pain. Concurrent associations of symptoms with anxiety and stress were weak or minimal.

Higher baseline anxiety, depression and illness worry, but not tendency to recall symptoms, were associated with stronger correlations between psychological states and physical symptoms. (Chapter 13)

Question 4 *Is there any evidence for consistent sequential relationships between symptoms over time?*

Hypotheses

Time series data from patients with MUS will not show predictable sequential changes to suggest psychosomatic causal sequences. This null hypothesis was refuted: sequential effects suggesting bidirectional influence between mind and body were shown using granger causality at a level which was statistically significant for data pooled between individuals. The data suggest heterogeneity of effects whereby visceral pain has a more powerful influence on subsequent emotions, while concern and stress are more potent precursors of physical symptoms. (Chapter 14)

Time series data will show an increase in concern and / or symptoms prior to GP consultation. In this group of relatively stable patients there were only a small number of consultations which were preceded by a rise in illness concern, while concern may play an important role in new symptoms(89), it appears to do so less in established ones. (Chapter 15)

Question 5: *Do symptom time series data for patients with MUS show signs of loss of complexity?*

Hypothesis

In keeping with Goldberger's model of adaptability as health(195), data from patients with persistently unexplained symptoms will show reduced statistical complexity. This hypothesis was confirmed: pooled data analysed using a bootstrap method demonstrated significantly reduced sample entropy, indicating loss of complexity (Chapter 16)

Question 6: *How do patients with MUS describe their condition after viewing their own diary data?*

Hypothesis

Patients with MUS will find the electronic diary, and report of their data, helpful in describing or understanding their experience. Participants were positive about the diary and, in many cases, about the findings from it at the feedback interview. Superficial qualitative analysis of individuals' narratives and comments during the post-diary interview suggested recurring themes of the unpredictability of their physical symptoms, the explicability of their mood and emotions as a secondary

effect to their symptoms and their self-portrayal as strong individuals, inclined to a defiant response to adversity. (Chapter 17)

Question 7: What does it mean to consider MUS as a complex illness?

This will be discussed further in Chapter 21.

Chapter 19 Discussion – strengths and limitations of the study

Study Sample

Sample size

The sample size (26 participants who completed diaries) was modest but comparable with other studies. While some electronic diary studies have used sample sizes of 80-100(115) these have either been for shorter studies or had a more flexible data entry regime(152). Longer studies such as the current research have used smaller samples: for instance a one year study had 24 participants(178) and a 12 week study of fibromyalgia patients used 14, albeit as a pilot. A sample size of 30 patients with temporomandibular pain was sufficient for conclusions to be drawn statistically(163). Studies of sequential change of symptoms using paper diaries have all used smaller numbers than the current research(49;140;238;239).

As this was an exploratory study no a priori calculations of statistical power were made, and decisions on sample size were made on the basis of prior published research and limits of practical feasibility. The original aim was for 50 subjects with two or more functional symptoms, identified by their GP at the stage of completed investigations, but scheduled review after the first 5 participants showed that this could not be achieved. A revised schedule, accepting ongoing rather than recently investigated symptoms and a smaller sample size was drawn up with a target of approximately 30 subjects.

Recruitment difficulty

Recruitment caused more problems than anticipated. With medically unexplained symptoms reckoned to account for around a fifth of GP consultations(17) one might expect no problem, however Verhaak has recently demonstrated that persistent MUS are less common(65) with an estimated prevalence of 2.5% in the entire population and approximately 1.5% of the population aged 20-64. This figure is similar to the 1% of the population of five Edinburgh general practices who were repeatedly referred to hospital for medically unexplained symptoms(240).

Even so, there were few referrals from GPs despite their initial enthusiasm and many of those were patients, not with typical MUS syndromes such as IBS and fibromyalgia, but with patterns of symptoms that the GPs could not adequately explain at all. Some appeared to represent the “heartsink” patients of earlier reports(241). There are several possible explanations for this referral of small numbers of more severely affected patients. Firstly it may be that GPs believe that syndrome labels such as IBS represent discrete conditions which do not require a somatoform or “functional disorder” approach. Secondly they may acknowledge the limitations of the symptom syndromes but take issue with the semantics of the MUS label, such that symptom syndromes are seen as “adequately explained” rather than “unexplained”; this would explain their referral of those one might term “medically inexplicable patients”. Thirdly there may have been a difficulty in acknowledging uncertainty – to offer to refer a patient with unexplained symptoms implies that the doctor is powerless to explain what is going on.

None of these options was explored with GPs, however several self-referred study participants volunteered that their GP had in some way acknowledged their powerlessness, although generally this had been by withdrawing from the problem rather than supporting the patient through their uncertainty. This difficulty in referring may have parallels with the work on GPs’ actions in consultations in which doctors appeared to misinterpret requests for emotional support by patients as requests for certainty which they as doctors were unable to provide(41).

Participants who self-referred themselves to the study appeared to show little doubt that “medically unexplained” was an appropriate term for their symptoms, even when they had a symptom syndrome label. It appears that while, to doctors, medically unexplained means “unexplained by medical disease”, for patients it means “unexplained by medical practitioners”. Several of these participants had stopped going to their GP because it was not helpful.

Illness categories

This study used broadly defined illness categories. Entry to the study did not require any physical examination, independent confirmation of diagnosis or a requirement to meet criteria on any symptom score or psychological test. This was intentional, with a view to obtaining data from typical patients in primary care, rather than a highly

selected group, hence the simple entry criteria of physical symptoms which could not be adequately explained by organic disease. Similar pragmatic criteria were used in studies of prevalence(17) and more detailed exploration of consultations(242).

Recent trials of interventions for MUS in primary care have also used relatively weak entry criteria such as screening for 4 or more symptoms in a checklist(133) or a requirement to have two or more unspecified symptom disorders(243). Given the limitations of current diagnostic categories(244) this strategy is appropriate, particularly for a naturalistic study.

Illness Severity

Reviewing the results of this study it is important to recognise that they represent a group of patients many of whom had high levels of physical symptoms, psychological distress and social disability as indicated by baseline questionnaires and personal biographical details in interviews. In comparison, treatment trials in primary care have focused on larger numbers of less severely affected patients: for instance in trials of reattribution, Larisch(133) screened 847 routine GP consulters, identified 319 for further interview and recruited 149 trial participants (17.5% of the original sample) while Blankenstein(132) identified 10 patients per GP with frequent attendance and five or more somatic symptoms. The participants in this study are probably more analogous to those seen in, for instance, the tertiary clinic population studied by Salmon(40).

Diaries

Compliance

Compliance with data entry was good (88%) and comparable to that seen in three diary studies of similar duration(152;153;172) which recorded between 80 and 90% compliance. The current study was unusual in using twice daily entry and in order to maintain this diurnal structure, some data were discarded from days on which only one entry was made. Use of the stricter paired data entry rate reduced the compliance rate to 80% which was still acceptable.

Internal validation

Internal validity checking by randomly presenting one of the visual analogue scales a second time during a data entry session showed almost all subjects and all variables

had excellent reliability based on the intra-class correlation coefficient. This measure is not widely reported within diary studies, and may not often be asked. It is technically feasible however and should be integrated into future electronic diary studies. The measures of compliance validity were used to weight data during the meta-analysis, thus adding to the reliability of this method.

Reactivity

The issue of reactivity – a trend due to the use of the diary - is impossible to separate from an independent change in the condition, for example due to resolution of the problem but nonetheless requires to be considered. Studies have shown, certainly in timescales up to a month, that reactivity is not a problem with electronic diaries(158;161). In the current research there was a significant trend in the mean, and also reduction in variance, over time. However some participants clearly reported resolution of their symptoms during the period of the study so this, rather than reactivity to the diary may have accounted for some of the trend. Concern about symptoms appeared to show greater downward trend and reduced variance over time than other measures and it appears that there was more reactivity with this than the other measures. .

Appropriateness of the visual analogue scales

Despite the plethora of complex rating scales for psychological constructs and illnesses, there is good evidence that single questions have considerable validity for mood(245). While epidemiological and intervention studies have used detailed question sets(133;246), a decision was made to use simple measures, firstly to aid compliance, and secondly because of the nature of multivariate analysis which meant that too many variables would be counter-productive. It is worth noting that the absence of specific psychological measures was a feature of the recent exploratory primary care MUST trial(136).

Participants seemed able to answer the questions comfortably and there was no evidence from the validity checking that one question was less reliable than another, or that there were a significant number of reversals (where high and low values are transposed), although this could not be excluded in all cases. The diary was designed to minimise this risk by keeping the orientation of each scale consistent for all recordings. Over-familiarity was avoided by randomising the order in which

questions were presented, and monitoring of the order in which questions were presented confirmed that this took place.

Diaries and assisted recall

During the follow up interview, participants were shown a time series chart of their symptoms and asked if there were any times or events that they recognised from the graphs. Usually there were none, however one participant immediately identified an episode of severe pain and anxiety and was able to use this as a trigger to graphically describe her catastrophising reaction to a flare up of pain. This anecdote suggests that while not a common phenomenon, the specific enquiry into ratings from one occasion can trigger relevant discussion.

In contrast to the method in this trial where data were collected for later analysis and associations generated statistically, Blankenstein, in a trial of reattribution in primary care(132), used paper diaries in an attempt to encourage participants to link symptoms and emotions at the time. Of 51 patients who entered the active stage of the trial one or more psychosocial links was made by 23, but only 8 related symptoms to depression and 2 to an anxiety disorder.

Diary ratings as a reflection of actual state

While the validity testing showed that participants' entries were reliably reproducible, there was no way of knowing whether the entries were a reliable reflection of the participant's inner state. This was pointed out by one participant who reported that her teenage daughter had watched her marking the "mood" scale and had commented to her that the level she expressed was not the level that everyone else would have reported. With hindsight this participant felt that her daughter had been correct, but – and this sentiment was shared with others – the diary had made her more aware of her emotions on a daily basis; this awareness was seen as constructive.

Analysis

Missing data

The study and analysis were designed to be robust to missing data. It was recognised that data entries would be missed and no attempt was made to impute these to construct complete time series. The decision not to impute missing data was made on

the basis of the marked day to day variability in many series which meant that prediction of missing values would be difficult. No formal simulation of missing values was carried out to test this further. The decision not to impute missing data had the effect of reducing the availability of more structured time series analytic techniques such as ARIMA models. Where only one reading for a day was available this was also excluded from the analysis allowing the twice daily structure of the data to be maintained. Discarding this data may have reduced the sensitivity of the analysis and increased the risk of missing a real effect (type 2 error) however it was hypothesised that if the within-day structure was maintained then any diurnal effects should balance each other out equally, particularly in the sequential analysis thus reducing the risk of introducing false positive findings (type 1 error).

The missing data could also have adversely affected the sample entropy analysis of complexity, however this has been shown to be robust to missing values(247).

Autocorrelation

Autocorrelation was clearly demonstrated within the data, as expected. This was effectively removed by the differencing process so differenced data was used for calculation of concurrent correlation coefficients.

Meta-analysis

The meta-analytic combination of individual patient correlation coefficients has not been reported elsewhere, however meta-analysis of correlation coefficients is a widely recognised technique within the social sciences(222). In the current research, each within-patient study can be seen as a separate experiment and matching variable pairs were combined using a random effects model which can deal with heterogeneity(188;222). Following Hunter & Schmidt(222) no statistical tests for heterogeneity were carried out as they argue that these can sometimes produce spurious negative results and should not be used when the distribution is wide. Whitener supports this view, recommending demonstration of credibility intervals(224) which should not be greatly wider than the meta-analytic confidence intervals if the sample is heterogeneous.

The random effects model of meta-analysis bears similarities to multi-level modelling in which analysis is nested at the within- and then between- subject level.

Meta-analysis was chosen over multilevel modelling for this exploratory study because the aim was as much to demonstrate the heterogeneity between individuals (as achieved by the forest plots – a typical output of meta-analysis) rather than identify mediating factors and develop “best fit” regression models.

Analysis of interviews

The interview reports in this research represent only a superficial inductive assessment of participants’ comments as recorded in written field notes.

Interpretation was carried out later, after all reports were available, sometimes leaving a lag of over 18 months between the original notes and the analysis. While this meant that there was no formal opportunity to develop and test hypotheses in later interviews, it is likely that this happened at an informal level during the data collection phase.

The key aim of this element of the research was not to develop new theory, but to confirm and elaborate previously published concepts within the current study population with a view to developing therapeutic interventions in the future.

Chapter 20 Discussion – anxiety, depression, concern and functional physical symptoms

This chapter reviews a wide ranging literature relating mood and cognitions to medically unexplained physical symptoms with reference to the current study. It deals with mood separately from anxiety and concern, although inevitably there are overlaps and in each section compares current findings with prior studies then elaborates possible explanatory links and models.

Current mood – the effect of low mood and depression

Findings from the current study

This study found moderate correlations between negative changes in mood and physical symptoms, particularly painful ones. For instance the meta-analytic correlation coefficient of lowered mood with muscle pain was 0.36 (95%CI 0.29-0.42) and with headache was 0.31(0.14-0.48). These values are conservative as they represent changes, obtained from differenced data, rather than the correlations of actual values which, while higher, were shown to be susceptible to the influence of autocorrelation.

When participants were split into two equal sized groups according to HAD-D at baseline, those in the higher scoring group (score > 6) all reported either muscle pain (11 participants) or headache (5) as one of their chosen symptoms. This HAD-D threshold is close to that of >7 identified as the optimal threshold for case-finding of depression in primary care in a recent large community based survey(248). While in this study no formal diagnosis of depression was made, it is likely that several participants would meet research criteria for depression.

When the results of the meta-analytic correlations were inspected in detail, higher correlations between changes in mood and symptom were seen in participants with higher HAD scores for headache and joint pain. This difference was less clear cut for muscular pain, but this analysis included only 4 of 15 subjects from the low HAD-D group. Thus it appears that higher baseline depression is associated with greater correlation between changes in mood and somatic pain symptoms. In contrast with

somatic pain and headache, visceral pain and other symptoms were less strongly concurrently correlated with depression.

Analysis of the sequential data suggested that the interactions involving mood were broadly balanced with equal overall weight in the psycho-somatic and somato-psychic direction. Low mood appeared to more strongly influence subsequent fatigue than other physical symptoms and to be influenced more by prior visceral pain.

Associations between depression and specific syndromes:

Chronic Pain

Pain and depression commonly overlap both in patients primarily presenting with chronic pain(249) and with depression(250). This relationship appears both common and complex, with interdependence of mood and symptoms going beyond the simple emotional consequences of either the physical suffering or social disability of pain.

The sequential study of depression and pain has yielded results suggesting that pain is a risk factor for subsequent depression, but also that depression is a risk factor for subsequent pain. This applies both to functional pain syndromes such as tension type headache(251), and the more “organic” patterns of neck or back pain(252).

Other syndromes

Depression is common in patients with other functional syndromes, including fibromyalgia(253), irritable bowel syndrome and chronic fatigue syndrome(254). A recent meta-analytic review demonstrated highly significant associations between depression and four different functional somatic syndromes(255). The effect size statistic d representing the strength of the association between depression and each symptom syndrome compared to healthy controls was in the range 0.62 -0.83 for fibromyalgia, functional dyspepsia or IBS and higher for CFS (1.34). These between subject effect size are larger than those seen between depression and functional physical symptoms in this study: d is approximately equivalent to $2r$ (222) which translates to values for d in the range 0.3 – 0.7 for physical symptoms and depression at the within subject level. Part of this reduced effect size is due to the use of differenced data to remove the effects of autocorrelation. Similar analysis with undifferenced paired data yielded higher values for r in most cases (for instance fatigue with low mood $r=0.55$, equivalent to $d=1.1$).

Co-variation in mood and symptoms

A small number of studies have measured mood and physical symptoms repeatedly using pen and paper diaries:

Healthy volunteers

Persson and Sjoberg(140)obtained twice daily self report data over 28 days from 8 healthy women. Symptom ratings showed strong autocorrelation but mood was essentially random. There were moderate correlations between symptoms and low mood, with Pearson's r ranging from 0.12 for simple skin irritation to 0.43 for "neurasthenic symptoms".

Brown and Moskowitz(256) used thrice daily self reports of symptoms and affect over 20 days, with affect and symptom scores based on participants' choice of verbal descriptors. They used a similar approach to this study by carrying out meta-analysis of all within-person analyses and demonstrated considerable heterogeneity. In the pooled results, however, current symptoms were more strongly predicted by prior symptoms ($\beta =0.41$) than by current unpleasant affect ($\beta=0.15$) and prior unpleasant affect ($\beta =0.04$).

Irritable Bowel Syndrome

While two studies of IBS have analysed stress or hassles(49;239), only one published single case study has analysed mood and IBS(257), in a patient with bipolar disorder. Correlations were non-significant after controlling for autocorrelation.

Fibromyalgia

Electronic diary series of fibromyalgia patients have consistently shown concurrent correlation between low mood and pain(156;166;258) without any evidence of a lagged effect. In their most recent study, in which the within person correlation between pain and low mood (after adjusting for autocorrelation) was $r=0.26$, Tennen, Affleck et al(258) also identified associations between prior depression and coping strategies, such that although subjects with a past history of depression did not experience greater pain, or lower mood during the study, they were less able to cope with painful days. A similar effect of prior depression on coping abilities was observed for patients with pain due to rheumatoid arthritis(259).

Using a different analytical process, and monitoring recovery during therapy, Feiler(148) identified a modest but significant coherence between depression and pain, however this was largely mediated by self efficacy. During recovery, the return of a sense of self efficacy was associated with improvements in daily pain and in mood

In a recent study in which diaries were used during a randomised controlled trial of pharmacological treatment for fibromyalgia, a greater day to day variation in pain was correlated with response to placebo but not to active treatment(260).

Temporomandibular dysfunction (TMD)

Using electronic diaries of pain, catastrophising response to pain, and mood (expressed as a choice of words), Litt(163) and colleagues identified weak but significant effects of mood on concurrent pain and on subsequent pain (high arousal negative affect state, $\beta = 0.13$ and 0.18 respectively). Again mood appeared to exert its effect through cognitive processes around coping, catastrophising and self efficacy.

Tension Type Headache

There have been no published diary studies relating tension type headache to cognitive variables and, despite the widely used name of “tension headache”, findings from epidemiological studies linking emotions to headaches have had mixed results. Only recently have methodologically rigorous laboratory experiments(261) demonstrated that headache prone individuals with depression were more likely to experience headache after stressful tasks than either headache-prone non-depressed subjects or normal controls.

Models for the link between depression and MUS

Epidemiology and shared risk factors

Several epidemiological studies have shown that depression and MUS (or somatisation) have overlapping risk factors, many begin early in life and include parental illness behaviour and personal experience of childhood illness (87;262), disordered relationships(263;264) and attachment style(265) and have led to the suggestion that these are sufficiently consistent that individuals at high risk(266) could perhaps be identified early in the course of their illness history.

A large literature, not reviewed in detail here, has demonstrated that experience of either depression or MUS increase the likelihood of an individual experiencing the other in future(91;267).

Insights from neuroscience

Neuro-science, particularly arising from applications of functional neuro-imaging, is beginning to show the intimate interaction of cognition, emotion and experience with incoming pain signals to the central nervous system(4;268).

Some of the overlap between depression and symptom syndromes, particularly those involving pain processing, appears to be explained by shared neural pathways, with overlap in neuronal activation between physical and social pain (269). Not all aspects of pain are closely neurally associated with depression: a functional magnetic resonance imaging study of induced pain in patients with fibromyalgia showed that depression influenced the affective component of pain but not the sensory component(270). This conceptualisation of pain as having separate affective and sensory components with their own neural systems is valuable and perhaps goes some way to explaining why troubled emotional processing influences pain(4). A further differentiation, between the effect of visceral and somatic external pain, is supported by a small study of patients with FM and IBS which showed greater activation of brain areas associated with emotional responses by visceral pain(271). Interestingly this finding is in keeping with the results of the current study in which visceral pain is associated with subsequent emotional distress.

At a cellular and molecular level, recent studies suggest that the glial cell network in the brain is not just a mechanical support structure but is immunologically active in supporting, or inhibiting neural function(62;272), with pain producing glial effects which then cause adaptations elsewhere.

Mood as a cognitive input

In parallel with neuroscientific links between mood and symptom perception and response, cognitive research is providing useful insights to the paradoxes presented by participants in the qualitative aspect of the current study.

Recently Vlaeyen and Morley(273) reviewed the evidence for an interaction between mood and the drive to achieve goals, such as completing a task or behaving a certain

way despite difficulty, using “stopping rules”. These constructs are cognitive rules about when it is appropriate to cease a task and include the principle of completing “as many as can” (AMAC) or its opposite “feel like discontinuing” (FLDC). They identified a model whereby the effect of mood on achievement depended on the cognitive stopping rule in place at the time. In the presence of negative mood they postulated that the AMAC rule would lead to greater achievement while positive mood would permit a breach of this rule and an earlier stop. On the other hand negative mood in the presence of the FLDC stop rule will lead to task avoidance and an earlier stop, whereas positive mood will overrule the stop rule.

This stopping rule model is one way of accounting for the paradox of why pain leads some individuals to avoid activity and become deconditioned (as for instance appears to happen to patients with back pain who demonstrate pain related fear), while it drives others to carry on despite stop signals in way which resonates with the “all or nothing” behaviour pattern as has been shown in fibromyalgia(274). The model provides a theoretical explanation for the behaviour described by several participants in this study, particularly around the onset of the condition, when the worse they felt, the more they pushed themselves to do more. Interestingly one, who exemplified this approach, volunteered that she knew at the time she could not be depressed because (from her experience) depressed people gave in and couldn’t be bothered.

Vlaeyen and Morley further speculate that therapy may be effective by changing stopping rules, for instance with “pacing” encouraging the FLDC rule in place of AMAC for patients with fibromyalgia. Acceptance (275) may also be construed as a reframing of stopping rules, perhaps implementing a different rule set geared to adaptation and flexibility.

Although this study did not examine the concept of stopping rules, the idea of “all or nothing” behaviour, which is conceptually similar to the AMAC rule, rang true for most of those participants who were asked. Several gave clear descriptions of pre-morbid effective coping styles under pressure in which additional stress led to better performance. The idea that negative mood, perhaps because of a diminishing satisfaction or because of other difficulties might drive the person harder before

becoming ill (overloaded being one acceptable metaphor) seemed an acceptable way of coming to terms with illness and adapting to it.

Summary

There is strong evidence from the current study that simultaneous changes in mood and painful functional symptoms are linked within subjects. This matches the epidemiological evidence for concurrence of MUS and depression between subjects listed in Table 2. This appears to be bidirectional, with an effect size of between $r=0.15$ and $r=0.35$. This can be rephrased as showing that one accounts for between 2% and 12% of the variance of the other, independently of prior levels of the that dependent variable.

What the current research finds in relation to mood and MUS is that sequential effects between mood and symptoms are relatively weak, but that two particular associations: low mood increasing future fatigue and visceral pain reducing future mood seem more prominent. In general however the effect of mood seen in the study is in a model in which emotion sets the context within which events, including unpleasant physical sensations are both interpreted and reacted to. Research elsewhere suggests that this is through shared neural systems and intermediary cognitive processes such as coping, catastrophisation, and stop rules.

Furthermore it appears that much of the biology and cognition of depression and chronic pain appears to be an evolutionary adaptive state to continuing injury, infection, or social stress. As such, patients with depression and pain will often have concrete examples of how their (mal)adaptive strategies have been effective in the past. By viewing pain and depression as overlapping adaptive phenomena, it becomes important not to see one as causing the other, but to see both as happening together in the same context.

Stress, anxiety and concern

Typically patients with MUS have been perceived as anxious and worrying excessively about their condition(72;73). A relatively small groups of patients with MUS meet the criteria for a full diagnosis of hypochondriasis(24), which is increasingly viewed as a specific disorder of health anxiety(244). While there is evidence that health anxiety between individuals correlates with healthcare

attendance(100) the effect of this anxiety, at least in primary care, is that physicians investigate and treat more often, even though patients do not appear to directly wish it(276). While general anxiety appears to influence overall tendency to consult, specific anxieties and concerns, such as fear either of the suffering of pain, or that ignoring pain will lead to worse harm, are prevalent in some conditions, for instance chronic back pain(273), they appear less important in more widespread pain disorders such as fibromyalgia(274). Current research into the neurophysiology of anxiety and the autonomic nervous system shows promise for MUS, particularly the functional gastrointestinal disorders. Both afferent and efferent vagal transmission are important in interoception and its effects on cognition and emotion (277) and studies of autonomic balance consistently show relative underactivity of the parasympathetic tone in IBS (278).

The study set out to test the hypothesis that anxiety and illness concern would be related to symptoms and that they should increase and decrease in relation to consultation with a GP.

Correlations between anxiety, stress, concern and symptoms

Studies of the association of anxiety with physical symptoms and its relationship to the decision to consult have yielded conflicting results. Major life stress (such as bereavement or unemployment) appears to be a stronger predictor of distress and consultation than accumulated daily hassles. Specific measures of health anxiety, illness behaviour and recent major life events independently predict new MUS(78) and subsequent GP consultation(279). This is also seen in the subgroup of patients who develop abdominal pain typical of IBS in whom psychological distress, and high ratings on scales of health anxiety illness behaviour all independently increase the risk of future abdominal pain(280).

Influences in the past appear to be important in sensitising individuals to stressful life events. Prior abuse is common in some groups of patients with MUS(281;282), but more subtle social interactions, for instance within the mother-child dyad also appear relevant(283). Lack of current social support both increases susceptibility to pain(284;285) and the likelihood of consulting(286).

There have been few studies prospectively tracking health anxiety and perceived stress or daily hassles and physical symptoms. While some have suggested associations between stress through daily hassles and symptoms(284), others have not. Diary studies suggest that in Crohn's disease(287), Gastro-oesophageal reflux(288) and IBS(49)there is an association between daily hassles and symptoms. Only in the case of IBS can this be detected over any time scale other than the concurrent measure.

Four studies have analysed variation in IBS symptoms and included anxiety or stress, all used pencil and paper diaries:

Suls(239) collected end of day IBS symptom ratings and most troublesome daily event data daily over 3 weeks in 44 subjects. Pooled data suggested no overall association between stress and symptoms though within subject correlations varied markedly with wide confidence intervals.

Stevens(289) collected end of day IBS symptoms over 8 weeks in 25 subjects with formal measures of anxiety and depression at baseline only. Analysis showed that symptoms were autocorrelated with clusters of high symptom levels. Symptom load was non-significantly correlated with baseline psychological state.

A pilot study by Dancey(238)used daily IBS symptoms and weekly hassles checklists in 30 subjects and found that symptoms in one week predicted hassles in the next but hassles did not predict symptoms.

A more detailed study from the same group(49)collected end of day IBS symptoms and a detailed list of hassles from that day in 31 subjects. Like Suls(239) and the current research they used an idiographic approach to analysis and found heterogeneity between individuals. Concurrent associations between same day stress and symptoms were significantly positive for only 14, even though this was without adjusting for autocorrelation which, as in the current study, was shown to be present. Using lagged regression of one variable against the other, Dancey identified significant effects of stress on symptoms from between 2 and 4 days previously on symptoms in 13 (42%) subjects and the converse, with symptoms influencing stress over similar time periods, in 11(35%).

This second study by Dancey(49) is the closest in structure to the current research, although confined to IBS symptoms (scored between 1 and 7) and daily hassles. While the hassles measure was a list of 117 possible hassles (all scored between 1 and 4), the analysis was carried out with a single hassles measure, the sum of all 117 scores, which may be rather more analogous to the single VAS used in the current research. While Dancey showed lagged effects of an independent variable, she did not account for lagged values of the dependent one in the granger causality model. Thus, Dancey's lagged effects may all have been artefacts of autocorrelations within the two series.

In the current research the concurrent effects of anxiety and stress on symptoms were less than that of depression. Indeed concurrent correlations for stress were often not statistically significant at all. Perhaps because of the stage of their illness, where most patients accepted that their symptoms were medically unexplained, few presented high levels of anxiety about the diagnosis and most tended to play anxiety down in interview. The exception to this was the more specific concern about symptoms, which in some cases, such as abdominal pain, was strongly concurrently correlated with the symptom itself.

Nonetheless, the pooled granger causality data suggest that subtle sequential effects were present and that these appeared more important in some instances than others, for instance the effect of internal pain on concern and the understandable association between anxiety and autonomically mediated symptoms. Interestingly, stress appeared to behave differently from the other three psychological variables in that it appeared to be relatively more associated with increased subsequent external musculoskeletal pain but less with fatigue.

Stress, concern and the consultation

There are no electronic diary studies relating to anxiety or illness concern and the decision to consult a doctor. A weekly phone interview study by Cameron et al(89), testing a model for choosing to consult derived from Leventhal's self regulation theory, showed that new physical symptoms which were ambiguous (i.e. they may have been due to physical disease or simply temporary changes in bodily function) were tolerated if they occurred during acute life stress and were thus attributed to the stress. If they and the stress persisted beyond around 3 weeks, they were then more

likely to be attributed to possible illness. The authors argued that in the acute stage of life stress, symptoms were absorbed as another element of the emotional impact but as time went on, the burden of accumulated emotional and physical load led to consultation.

While work such as that of Cameron has viewed illness concern from an attributional perspective(89), and the model of somato-sensory amplification whereby anxiety causes physiological effects which are amplified by hypervigilance(217) feeds well into this, earlier work by Zola(7)suggested that reasons for consultation were more specific and immediate and relevant because of the interruption of social functioning, and that it was when symptoms became a social threat rather than when they reached some level of severity or changed attribution that people became patients. His 30 year old observation that failure to pay attention to the specific trigger within the consultation led to poor outcomes rings true today, but has not been empirically tested except as part of a more general personal care.

From the diary data and interviews, there was only one clear instance of overwhelming anxiety about a flare up of symptoms leading to catastrophic worries and an urgent appointment with the GP. This occurred in a woman for whom experience of illness in both her parents had graphically taught her that normalisation could be dangerous.

Anxiety within the consultation

Although most work has looked at anxiety as a risk factor for symptoms or a trigger for consultation, recent work has observed the effect of trait anxiety in the consultation(290) and shown that subjects with high levels of anxiety were more dependent on the physician eliciting biomedical information and less satisfied with consultations which had a high emotional content. Dealing with anxiety and events in the consultation also appears to depend on skills in emotional processing, for instance there are differences between fibromyalgia patients with and without alexithymia(291).

Part of the problem with anxiety in consultations is that anxiety about missing a physical illness is an appropriate adaptive response in doctors, particularly in general practice where a degree of uncertainty is almost always present and must be

tolerated. Therefore “unexplained symptoms”, which by definition have the potential to be due to hidden serious disease, should be anxiogenic to doctors. Appropriate adaptive strategies for the clinician may include processes such as hypervigilance and reassurance seeking which lead to further investigation and referral. As if this were not in itself unhelpful, it is plausible that patients are then perceived as anxious themselves through the unconscious defensive projection(292) of internal physician uncertainty onto the patient, whereby “I don’t like my own anxiety engendered by my diagnostic uncertainty” translates into “this patient is anxious about their diagnosis”. This doctor-perceived anxiety to do something, which is at odds with patients’ wish to validate the interaction between their emotional and somatic experience has been clearly demonstrated elsewhere(293).

Summary

Relatively few participants in the study described pervasively raised anxiety levels and for most, descriptions of anxiety were in relation to social function and ability to cope rather than worry about illness. Furthermore patients repeatedly reported occasions in which symptoms occurred when they were not anxious. Although the hypervigilance and somato-sensory amplification models are theoretically robust they were the exception rather than the rule in this group of patients and, may play a modest role in most persistent MUS seen in practice.

While recent literature has stressed the personal cognitive aspects of anxiety in triggering medical help-seeking, older work and recent narrative studies suggest that social processes and the threat to social identity and function are at least as important. In this context, the quality of the physician patient relationship as a secure environment in which emotions can be acknowledged, but then put away again, may be as important as the content.

While stress has a role in the broader picture of symptom causation and consultation request, its concurrent association with symptoms appears too weak and inconsistent to use in mechanisms for patient explanations in most cases. However the greater effect of stress and anxiety on future physical symptoms shown in Chapter 14 may be more fruitful, particularly as explanatory models can then become less about reattributing current perceptions (caricatured as “you feel ill but actually you are anxious”) and more about how reactive and adaptive temporal patterns affect health

(for example “because you have been dealing with a difficult situation, your body has become unwell”). This change from correction to interpretation offers the potential for more positive interventions, whether at the level of normalisation in a first consultation, or of more complex work in the more established case.

Chapter 21 Discussion – variation, causality and meaning

Observation and Causality

The key results of this study are that the day to day variation of symptoms displayed temporal structure and that this occurred in different ways: within symptoms (autocorrelation and reduced entropy); concurrently between symptoms (correlations); and sequentially between symptoms (Granger analysis). Nonetheless there was no single predominant pattern common to all or any subgroup. Rather there was a continuum of effects, which displayed heterogeneity but not distinct cut-offs and differences.

Such empirical findings do not lend themselves to straightforward explanation, either by patient or doctor. While patients with MUS have no doubt they are ill, they are thwarted by their physicians' inability to give good explanations. Names or labels may help, although for many participants in the current research these were seen as weak and were applied diffidently, especially initially, although for some they subsequently provided a solid base from which to rationalise ability and disability. These reports matched earlier published findings(33;234).

In particular, blunt causal statements were regarded negatively: “(he) said it was caused by stress” was typically recounted in a contemptuous tone. Few participants described attempts by doctors to elaborate more complex mechanisms, although it is clear that lay models of ill health and symptom attributions are multiple and complex(34;51;294) and that doctors feel uncomfortable with unresolved diagnoses in patients with MUS. Clearly the notion of causality for the patient experiencing MUS is a challenge

“Multiply explained symptoms”

Historically, biomedicine has flourished within a positivist world view, perhaps best exemplified by Koch's postulates for the demonstration of a micro-organism as the cause of disease. These depend on specific tests to demonstrate successional causality in which a pathogen A precedes, and is a necessary cause for, illness B. The language and conceptualisations of this single cause approach spill over into much more complex explanations derived from epidemiological evidence but these are still

frequently simplified for practical use, for example with the cardiologist (or patient's wife) asserting that "your heart attack was caused by smoking".

Studies of illness narratives suggest that individual patients have their own explanatory structures, to which a rational deductive approach to making sense of cause is still usually applied: hence Kleinman's explanatory model for illnesses (232) structures patients' constructs of their illness as having a name, cause, timescale and other characteristics. Two studies from Montreal(295;296) have examined illness narratives of patients with MUS and identified a greater diversity of knowledge structures used to describe and explain illness which are listed in Table 25. This work emphasises earlier suggestions that deductive rationality is not the only method by which individuals make sense of their illness experience(297) and indeed for complex and unpredictable illnesses it may be necessary for explanatory models to simultaneously account for conflicting observations.

Table 25 Classification of explanatory structures

Class	Subgroup	Example
Rational Explanations (deductive models which may adopt scientific discourse or terms)		
	Causal factors	Exercise makes me feel worse
	Mechanisms	When I exercise the lactic acid builds up and makes me sore
	Labels	Exercise makes my CFS worse
Chain Complexes (temporally but not causally related phenomena)		
	Single	It happened when...
	Repeated	Sometime when... then..
Prototypes (other experiences or characteristics which are seen to share properties)		
	Self identified	Everyone in my family....
	Imposed	My doctor says people like this never...

Narrative studies also demonstrate that explanations have a purpose and a context. While doctors simply seek to diagnose and treat, patients require meanings for a number of purposes ranging from reduction of personal uncertainty to the negotiation

of social roles. Hence while biomedicine may value scientific validity above all else, patients need to find explanations and meanings which are “good enough” to fulfil a range of functions (298).

While participants in this study were still seeking explanations, they appeared to simultaneously hold and test multiple explanations, recognising that no single explanation could fit their situation. Despite their own multiple explanations, they reported how their own doctors rejected their proffered possibilities, instead seeking additional unseen factors of their own (such as stress). Through this negotiation of explanations, doctors were perceived as unable to articulate uncertainty except by acknowledging defeat.

Social construction of meaning in MUS

While the biomedical search for underlying truths and casual relationships, through epidemiological explanation continues, with no sign of resolution(244), an alternative is to take a constructivist approach whereby “reality” is socially constructed. This adopts a generative causality in which events and meaning are formed through the human interactions which serve to communicate experience, distress and needs within a social framework. Thus, interpretation of causality needs to consider the purpose of the narrative rather than view it as objectively independent. For instance a narrative analysis of women with chronic musculoskeletal pain(27) highlighted elements of plot (“how I became ill”), performance (“how I show strength in coping with it”) and argument (“why you must treat me with respect”) which only make sense in a social context, rather than as an internal personal or scientific factual explanation.

While a constructionist model brings insights into personal and social processes, it does so by abandoning generalisability. It also raises the challenge of agreeing meaning when social agendas may differ, for instance between patient, family, doctor and employer. Given that biomedicine is designed around a positivist framework of the concrete existence of disease, there are inevitable clashes between the realist perspective of science and this socially constructed worldview.

A critical realist approach

More recently there has been an attempt to synthesise these apparently contradictory world views in the epistemological concept of critical realism(299). Arising originally from Bhaskar's philosophical writing, this argues that both the "real" world of positivist science and the "social" world of constructivism co-exist and cannot be adequately be described by each other. The constructed emerges from the real and vice versa, but neither makes sense solely in terms of the other. However within both those worlds actions are described at three levels: the first is the empirical or experienced (which can be recalled narrative, or a scientific measurement – both replete with biases and error), the second is the actual (implying that there really is *something* but that it cannot be accurately known), and the third is the deep or mechanistic. This deep level is crucial to causation in the critical realist model because it is here that processes interact, in a way which is contingent on their environment. This causality is not successional, rather it is generative: actions depend on mechanisms operating in a particular context(28).

At first reading critical realism may seem like an attempt to restate common sense in a complicated fashion, but its advocates argue not just for its importance in resolving the otherwise conveniently ignored tension between positivist and constructivist methods, but also for its implications for research in general(28;300). In particular there are two arguments relevant to research in MUS: that subgroups and individuals are important and that similarities across research domains may be more valuable than differences within them. A third, newer argument, which I shall address in the next chapter, is that critical realism may overlap with, and possibly become statistically tractable through, complexity science.

The importance of subgroups

The critical realist argument for the importance of subgroups depends on the causal interactions between mechanisms and context such that some interactions will be more productive than others and that this effect is not simply due to random chance. In contrast, the dominant view in biomedical research conflicts with this: while subgroups may be interesting, study design should seek, through randomisation or stratification, to balance them and therefore cancel out their effects, leaving only the main effect under investigation. Meta-analysis in particular espouses this belief that

differences between samples must first be considered to be due to statistical variation before calling on real differences(222).

Commenting on their study of extended reattribution in primary care(133), Larisch and colleagues(134) carried out a secondary, exploratory, analysis of the heterogeneity of their study population which comprised GP attenders with multiple physical symptoms and at least mild emotional distress. They identified three adequately distinct clusters: predominantly psychological distress; predominantly physical distress and combined high physical and psychological distress. The two groups with high levels of psychological distress were more able to make psychosocial attributions for illness than the physical distress only group. During the intervention, health anxiety diminished in both groups with psychological distress, but rose in the physical distress only group. Paradoxically, although emotional distress rose during treatment, in this group with physical but not emotional distress, satisfaction also increased: the authors interpret this as a group of “deniers” who found facing their distress uncomfortable but ultimately helpful. This post hoc subgroup analysis has construct and face validity, but was only published after the main study results showed very modest benefit for the sample as a whole.

In the current research the idiographic nature of the data collection and analysis mitigated against much cluster analysis, however the meta-analysis of individual correlation coefficients displayed statistical heterogeneity, suggesting that they did not all come from the same population distribution. Similarly the results of the granger causality tests also suggest that different patterns occur in different individuals, which are individually relevant but difficult to demonstrate as a group effect.

The presence of functionally different subgroups becomes critically important in planning interventions for the diverse conditions represented as MUS. Where the effects of treatment are modest relative to the statistical “noise” of diversity, demonstration of efficacy or effectiveness through conventional trial methods becomes extremely difficult. On the other hand, studies which reduce heterogeneity by selecting very similar participants lose out on generalisability beyond that subgroup.

Similarities across contexts suggest mechanisms

The argument for identifying similar effects across widely differing research contexts as a source of mechanisms has its origins in social and health policy research(300). By looking for common mechanisms across groups it takes a complementary approach to the focus on subgroups. In symptoms research, a number of psychological processes appear in a wide range of contexts: examples include constructs such as self efficacy(148;163), catastrophisation(112;113;116;301) and symptom-related fear(110;164;274). These all appear as strong correlates of disability and distress in a variety of both organic and functional disorders. Common mechanisms may also be found in dynamic models such as the notion of stopping rules and inhibition(213;273).

Study of these mechanisms, and dynamic processes is currently leading to better understanding of functional disorders and their role in recovery(148), but also to specific psychological treatments, for instance dealing with catastrophisation(116).

Interviews within the current research identified common themes across different symptom experience, for instance the “need to understand in order to manage” approach to symptoms, and all or none behaviour. Such processes were not inherently pathological, rather they can be seen as appropriate adaptive mechanisms which had become mismatched to their context.

Based on a critical realist model, the combination of mechanisms and contexts make it possible to develop adaptable models which can link the experience of physical symptoms with, or without life stress or threat, to somatic physiology, emotional distress and social disability.

The overlap with complexity

Complexity was first described as the science of emergence, in which behaviour of a system as a whole emerged from the interactions of the agents comprising it, in a way which could not be understood by knowledge of the agents alone(190). The interacting agents can be viewed, from a critical realist position as generating, and responding to “causal” mechanisms within their own context.

From an analytical perspective, complex systems do not fit the rules of independence and normal distribution on which parametric statistical models depend. Rather they

produce distributions of events which show clustering and patterns in time or space which show similarities across different levels of scale. In complex systems, subgroups are not just due to chance or noise, but part of the underlying structure. Thus, a complex model is entirely compatible with the three levels of critical realism: the deep interaction of mechanism and context, from which emerges the actual system, on top of which sits the experienced and recalled version of this.

In the current research a robust statistical method for identifying loss of complexity was hypothesised, and showed significantly reduced complexity in the actual data compared to surrogate data sets. The metaphorical notion of a chaotic narrative met with recognition from several of the participants and even simply accepting that there were no simple explanations that could explain things appeared to be valuable. When participants saw their data plotted as a time series graph it had the effect of confirming the bewildering experience of trying to make sense of daily symptoms.

The symptom diary as a critical realist tool

The combination of the quantitative and idiographic makes electronic diaries such as those used in this study an unusual, but not unique, research instrument. Adding this to a reflective interview in which the patient's narrated experience and a doctor's interpretation of the data is made both at a personal level and within a general context acknowledges the multifaceted reality of the individual. The method seeks out combinations of contexts and mechanisms which are personally relevant and at the analysis and feedback stage makes them real, explicit and able to be viewed in a new light. Indeed there may be some similarities with the way in which patients give doctors the permission to reveal, through X-rays, that which cannot otherwise be seen(302).

Quantitative diaries may also have a therapeutic effect, through reflection and possibly through facilitating confidence in exposing one's inner concerns(182;303). This blurring of the difference between assessment and treatment is again part of the real world approach.

Implications of a critical realist approach to MUS

From a critical realist approach, possible models of causality in MUS extend far beyond the conventional successional causation of somatisation, which views

unhappiness as incorrectly expressed through physical distress. Instead, effects are seen as contingent on circumstances rather than consistent and reproducible. While this approach has the advantage of permitting flexible and personally relevant models for cause it introduces difficulties for describing and evaluating treatment.

The implication of accepting that illnesses arise from the contingent interaction of contexts and mechanisms is that any given treatment, while ideal for one patient, may be wholly inappropriate for others. This approach contrasts with the risk factor modification approach of much modern medicine in which, for example, individuals' blood pressure or cholesterol are lowered towards population-derived ideals on the assumption that risk according to these is continuously distributed and shifting risk will result in some, randomly determined, individuals benefiting. Indeed it is closer to the "individualised treatments" to which pharmacogenomic research aspires, in the hope of targeting critical processes, or genes, for an individual in order to give personally effective medication.

This lack of a "black box" effect, common to all patients, may go some way to explaining why treatment trials using reattribution, which seem effective for some patients, have modest results when applied in wider trials.

Rather, instead of a "black box", doctors and patients may be left with a "tool box", both for explanations and for therapy. In this, a variety of techniques become relevant but the goals are to find personally relevant explanations, and base treatments upon them. This aim becomes similar to that outlined by Engel who described combining the understanding of the illness and the person thus:

the interpersonal engagement required in the clinical realm rests on complementary and basic human needs, especially *the need to know and understand and the need to feel known and understood*.

The first, to know and understand, . . . is a dimension of being scientific; the second, to feel known and understood, is a dimension of caring and being cared for. Both may be seen as derivative and emergent from biological processes critical for survival . . . (304)

Taking the approach of understanding illness in terms of mechanisms and contexts, requires sensitive listening for, and testing of, possible mechanisms. As well as validating the embodied experience in patient narratives which link emotional life and bodily experience, this listening may lead to fruitful attempts to harness and

channel those links in agreeing generative explanations. However this requires sensitivity to the multiple purposes of knowledge, as identified in constructionist models, and recognition of the risks of mismatches between different types, and implications of, knowledge. However, if the starting point is that explanations have to be “good enough” rather than absolutely correct, the doctor becomes able to explore shared models with the patient, enabling both parties to bring their own experience and knowledge to solving the problem.

This focus on reaching understanding appears to suggest that reaching an imperfect explanation may be better than reaching none. Not only is this compatible with Salmon’s identification of the need for enabling explanations(40), but it also fits with Frosthalm’s finding that patients’ uncertainty about the nature of their health problem is a strong predictor of dissatisfaction(233). This importance of reaching *an* understanding rather than *the* understanding presents a real challenge to the biomedical system which reifies diagnoses. Kirmayer takes this view further, arguing that the application of authoritative diagnoses restricts the capacity of the individual and doctor to “improvise” useful meanings, and create possible actions for dealing with illness(305).

From this critical realist position, classifications become not concrete categories, but labels which justify action – be it intervention by a physician or acceptance of an individual by society. It becomes unnecessary for doctors to hand down simplistic explanations to patients, for they already have complex constructions of possible meaning and, at least from the experience of this study, value the doctor engaging with and interpreting that complexity.

Testing a critical realist model

To test this model, wherein apparent heterogeneity is real (rather than simply noise) and due to differences in underlying mechanisms, requires a change of emphasis. Instead of a generic treatment for somatisation, the focus may need to shift in four directions:

Firstly interventions may be targeted not at the overall problem but at putative mechanisms: for instance selecting those individuals in whom catastrophisation is

prominent, or patients with all or nothing thinking, and addressing this as well as the specific symptoms complaint.

Secondly a “toolbox not black box” approach may be tested, whereby the focus is on interpreting illness and revising functional goals rather than labelling and curing disease. This overlaps work around acceptance(275) and also the potential benefits of interpretive medical consultations(306).

Thirdly interventions may be targeted further “upstream” in the pathogenesis of complex problems by reconsidering the process of normalisation(42), whereby symptoms are constructively labelled and dealt with at an early stage to minimise uncertainty, and to permit the early expression of, and support for, the emotional consequences of troubled lives(307).

Fourthly interventions may need to be seen in the long term context. Thus instead of a short sharp treatment, maintenance strategies for living with uncertainty and dealing with exacerbations of symptoms within a longer term therapeutic relationship may be needed.

These possible trends, away from specific interventions towards more general process based models more dependent on the aetiopathological processes than the ultimate diagnosis, mirrors Luyten’s view of future trends in depression management(308) and represent a move away from the single diagnostic category model whereby all MUS are seen as expressions of a common process of somatisation of mental distress .

Conclusion

Conventional models of cause and effect do not fit complex disorders such as MUS. The perspective of a critical realist approach of interacting mechanisms and contexts offers a useful framework both for aetiopathogenic models of functional disorders and for individualised explanations. Unfortunately, in doing so it challenges the authority of conventional methods of gaining and testing knowledge, which depend on homogeneity and generalisability.

Chapter 22 Discussion - towards a model of MUS as complex adaptive illness

Building on the introduction to complex systems and non-linear dynamic models in Chapter 3, the analysis of entropy in Chapter 16 and the concept of critical realist assessment in Chapter 21, it is possible to argue that MUS are complex disorders, which can be understood from a complexity perspective. Complex systems are open to their environment as opposed to closed off, are constantly co-evolving with that environment and are directed by tendencies whose actual effects depend on their context. The real world patients with MUS whom this study investigated would certainly seem to fit this description and it is worth further exploration.

Which level is the system?

Viewing the patient with MUS it is possible to consider complexity at two levels: the within person system and the between person systems which comprise society or the healthcare system.

At the within person level, cognitive goals and strategies, recalled memories and expressed emotion all interact in the context of somatic and neurological physiological processes. These elements are densely interconnected and provide for a multitude of potential responses, which in turn confers adaptability to many different environments(4).

At the level of social interaction, a healthcare system can be viewed as a larger complex system in which consultation patterns emerge from all the individual interactions between patients, their doctors and society(14).

The emergent nature of complex systems in which behaviour at one level emerges from the interactions at a lower level means that both of these system descriptions are valid and both will be considered.

What advantages would complexity provide?

The nature of complex systems is that they react and evolve according to their environment. This capacity to evolve generally confers adaptability, such that as the inputs to the system change, the system's behaviour adapts. However in some cases adaptation and evolution can proceed so far, in order to thrive in one particular niche,

that there is a trade off in general adaptability to set against increased robustness within the particular context. In general, it appears that physiological systems, as they age or are affected by illness, lose adaptability and complexity(204).

This notion of physiological complexity, adaptability and health, has parallels in psychology and neuroscience with increasing evidence that a strong mental sense of coherence - enabling an individual to see meaning, purpose and opportunity in circumstances - is associated with ability to adapt to trauma(309) while in contrast the complexity or entropy of the EEG diminishes in patients with Alzheimer's Disease(208).

Implications of MUS as a complex disorder for clinical care

If one considers MUS as complex system disorders, there are a number of implications for both patients and doctors.

Indeterminacy

Complex systems which are maintained by multiple interacting processes over time, cannot be accurately predicted, even if the overall pattern of their behaviour fits a statistical distribution. By adopting a model with such inherent indeterminacy, doctors are forced to acknowledge the limitations of their knowledge, because full knowledge of the system is impossible. Acknowledging this indeterminacy, while at the same time recognising the overall system and its general behaviour, may be a way for doctors to be open about uncertainty, without relinquishing all authority or experience.

It is clear, both from the current study and from published work, that patients perceive their doctors as unwilling to admit the limitations of their knowledge. It is not, however, sufficient for the doctor to give in the diagnostic challenge and hand it back to the patient(40). Rather there is a need to find practical ways of dealing with this uncertainty within consultations.

Perpetuation rather than cause

While complex systems must have had a starting point at one time, their dynamic nature makes it more important to consider where they are at present rather than where they started. The model of a complex dynamic system rather than a single

illness with a trajectory from beginning to end may free doctor and patient from identifying exact cause, and allow them to concentrate on the present and future.

Driving processes

Along with the indeterminacy that the invoking of complexity as a model for MUS brings, comes the opportunity to focus on causal processes. A model of MUS as a complex illness will have physiological, neurobiological and cognitive elements and where necessary these can be explored. The focus shifts however from a single problem or label (such as IBS, depression or even somatisation) to a series of processes which may, but need not, be related in some way.

Many of these processes can be interpreted as adaptations, either successful or not, and their use can make sense as adaptations, or failed attempts to adapt. This focus on adaptive processes rather than flaws or weaknesses opens up possibilities for explanation which build on positive coping or adapting attributes rather than negative ones such as excessive worry or failure to cope with adversity.

Parallel explanations

Published work on illness attributions(294;310) and explanations(295;296) along with the interview data from the current study all point to patients having multiple explanatory models which are more or less formed, which have different purposes and which co-exist in a state of tension and ambiguity. Doctors too feel uncertain about explanations and at times are at odds with patients over beliefs and models. This tension is well described by Kirmayer who discussed the inherent tension between diagnostic interpretation (which is normative and concerned with classification) and therapeutic interpretation (which relies on the improvisation of meaning in order to answer the challenge of “how to continue”), concluding that:

The goal of patient and physician is to create enough certainty to diminish the threat of the inchoate while preserving enough ambiguity to allow for fresh improvisation(305)

The complex illness approach is inherently incomplete, describing context dependent processes not concrete and consistent effects and so does not need to insist on one dominant explanation. Rather it permits multiple models to coexist as different ways of viewing the system, even potentially when they are incompatible.

The danger of just allowing everything

Having stressed the openness of a complexity based model it is important to define some limits to prevent it simply becoming an easy catch-all for anything. However complex systems have their origins in strict quantitative science, and while it is easy to absorb the attractive metaphors into softer concepts, rigour is both desirable and possible.

What processes should a complex system have to account for functional symptoms?

The model of adaptive processes, evolved on the basis of the individual's past history offers several possibilities which can be elaborated. These are based on the premise that brain sensory systems respond to sensory input in a pre-conscious way using emotion dependent methods or matching inputs to recalled memories(4). Broadly speaking the adaptive processes in MUS form three groups: threat surveillance and escape; failure of suppression of continuing threat signals; and withdrawal and inhibition to avoid further damage. These ontological categories can then be expanded in terms of the physiological and cognitive processes which operate within them. The focus of these adaptive responses on threat goes back to the notion that symptoms are physical sensations that indicate that something may be wrong, whether internally or externally. The emphasis on threat is also central to agreed definitions of pain which include both unpleasantness and threat.

Threat surveillance and escape

This category includes both cognitive and physiological processes. Cognitive processes include surveillance and interpretation of risk and the related pathological construct of hypervigilance. Physiological processes are largely autonomic responses to arousal such as increased heart and respiratory rate and altered visceral function. It is already widely accepted that these processes play a part in MUS, and they may be applied across the whole range of conditions within reattribution models(46).

Evidence in this research suggests that anxiety has relatively little influence on many symptoms such as pain and fatigue however and it may be helpful to limit the threat surveillance and escape category to certain patients or symptoms where these predominate.

Failure of resistance

This category links closely with current understanding of chronic pain disorders. While the exact methods of neural transmission and suppression of nociceptive signals is incompletely understood it is clear that several different pain pathways and neurotransmitters(277) are involved and that inhibitory processes are in play at several levels from spine to cerebral cortex.

As the affective, or unpleasantness, characteristics of pain in some types of MUS appear to be transmitted separately from localisation information(270) and be projected more to brain areas associated with emotion(271) it seems plausible to consider some MUS symptoms primarily as pain disorders, wherein pain and emotion are closely interlinked.

Withdraw / repair

The endocrine and immunological processes seen in the fatigue associated with MUS, and particularly in chronic fatigue syndrome, appear to be different from those seen in depression, even though the conditions share many behavioural and cognitive features(254). Recent developments include recognition of subtle changes through hormonal (61) and neuro-imaging studies(311) which suggest a suppression of normal function. This may be interpreted ontologically as indicating a withdrawal state, akin to the cytokine induced sickness response(312) which may have been appropriate in the early stages of infection or psychological stress but then become firmly established. With increasing evidence that cytokines are in some way related to mood(312), even though the evidence for consistent immunological dysregulation in CFS is inconclusive(313), it is plausible that a withdraw / repair process should influence mood and characteristics such as perceived self-efficacy.

Interactions between physiology, emotions and coping

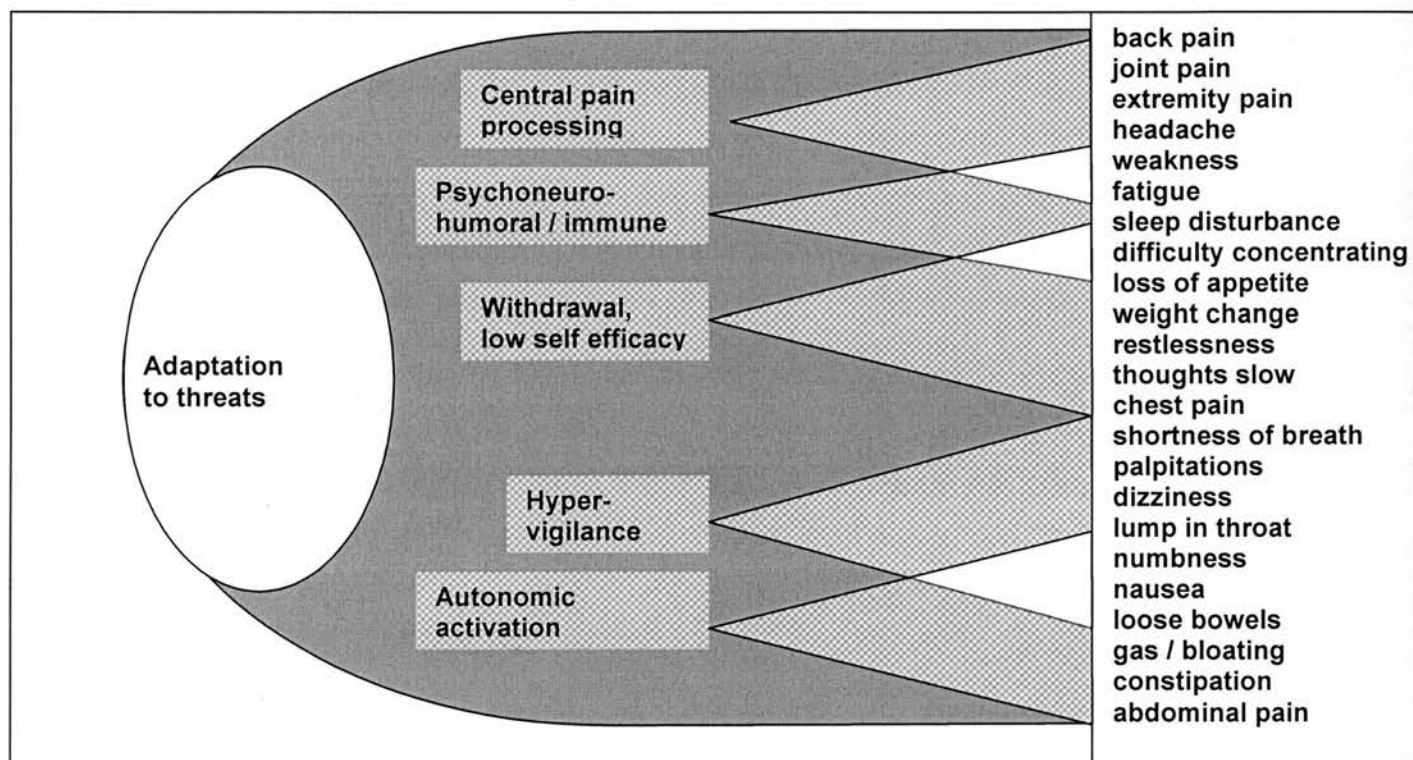
The three categories of threat adaptations which I propose – threat surveillance and escape, suppression and withdrawal – all have psychological and physiological components. In a complex systems model of MUS, symptoms, behaviours and methods of coping should emerge from these interacting processes. However as the pattern of emergence is contingent on both the current environment and past events, one should not expect the emergence of a particular pattern of symptoms or behaviour to be inevitable given a particular set of processes. Indeed as the

processes are reiterated, effective coping strategies such as maintained self-efficacy and sense of coherence, or dysfunctional mechanisms such as catastrophisation and emotional reassurance seeking may emerge from the threat adaptation processes and in turn may act as mediators between the emotional responses elicited by future sensory threat input and the cognitive and voluntary strategies used to deal with them.

A re-formulation of medically unexplained symptoms

Based on the above categorisation, Figure 31 shows a re-drawing of figure 1, this time replacing the overlapping syndrome labels with physiological and psychological processes. Instead of Deary's common factor(86) (which he argues is neuroticism) I have placed the common ontological goal of adaptation to threats.

Figure 31 revised model of MUS, after Deary and figure 1.



Implications of complexity for research into the management of MUS

Functional classification

If MUS are considered as complex illnesses which share some or all of a set of underlying processes, classification is likely to be difficult for several reasons. Firstly if the processes also occur in normal and organic illness behaviour it is difficult to classify them as pathological. Secondly classification implies stability over time, but it is clear that this is not the case, both from the variation described in this study over short time scales and the shifts in epidemiological studies over years(108). Thirdly any attempt at classification is dependent on the purpose of the classification and the tool used. Hence, depending on the perspective, functional abdominal pain can be seen as belonging to the same category as fibromyalgia (functional pain syndromes inconsistently associated with emotional distress and disordered central pain processing) or as separate (conditions affecting different body parts, with different illness concerns, causing different disabilities). Both perspectives are right, in the same way as a physicist may say red and green light are both similar waves from the visible electromagnetic spectrum with arbitrarily defined wavelengths while a driver at a set of traffic lights would interpret them as opposites.

Thus, for complex illnesses, classification is not absolute, but depends on the purpose of the classification.

Non-independence of characteristics

Typically in healthcare we deal statistically with relatively closed systems which are approximately normally distributed. Such distributions occur when values are independent: for instance the distribution of human height within a population is approximately normal, and the height of each individual is not (usually) affected by those around him.

Symptom counts, and associations between physical symptom counts and emotional symptoms are not normally distributed however. Several large epidemiological studies show that the distribution of unexplained symptom counts to be a smooth curve. From the data in published papers it is unclear whether these are closer to an exponential curve (as in half-life decay) or the power law seen in complex systems and in our study of consultation patterns(14). Certainly the curves show no obvious discontinuities. Non-independence is also evident in the associations of symptom

counts and psychiatric caseness, which, while arguably separate dimensions of illness experience, are not statistically independent – the prevalence of all psychiatric diagnoses increases with the number of unexplained, and explained, physical symptoms(68;70).

It is unclear at what level this non-independence occurs. It may be simply due to common factors, for instance childhood adversity, in which case it could be experimentally controlled for, however a complex systems approach would accept the non-independence as an inevitable consequence of the mutual interactions between different processes.

The implications of non-independence for research

Non-independence raises major issues for research into MUS, particularly when it comes to intervention trials. Two major issues will be considered: the contents of the “black box”, and the distribution of effect sizes.

Typically interventions, whether single drugs or complex and flexible psychosocial interventions are standardised. This is done in pilot work and leads to selection of the model which performs best in the pilot situation. In pharmacology it is reasonable to assume that the pilot context is similar to the main trial, but in complex interventions this may not be the case. Individuals and their doctors may behave very differently between different contexts and even within the same one. If one takes a complex adaptive approach it is reasonable to suppose that the interactions have evolved locally according to some adaptive gain, hence each separate context is the way it is, because of its history as much as its current measurable properties. This evolutionary dependence on context has major implications for the transferability of an intervention.

If one takes a complex systems approach to complex interventions, for instance the effect of changing the way a GP interacts with certain patients as in recent reattribution trials, then a number of additional problems arise, beyond that of developing an intervention which is appropriate to the individual contexts in which it is applied. These are to do with the way complex systems respond to environmental change. As described in Chapter 3, complex systems generally absorb changes with no major effect, but sometimes change more dramatically. This distribution of event

(or “effect”) sizes is not normal, but typically follows a power law distribution. This means that the majority of effects will be minimal, a few will be noticeable, and a very small proportion will be dramatic. General practitioners are used to this: much of the time we advise patients about lifestyle and even when there appears to be acceptance, little happens, but on other occasions unexpected and sustained changes occur as in the analogy of “tipping points”. Statistical analysis of trials is however designed to interpret responses as a range of responses representing a shift in the mean plus random error. This analytical framework is suitable for measuring efficacy of an intervention (i.e. whether it works) when all variables have been controlled for such that the effect is independent, and for evaluating effectiveness (whether it is a worthwhile investment) for a given healthcare situation. For complex interventions, particularly where the distribution of effect sizes may be non-normal such an approach retains its value for effectiveness (where a finite resource is available), but gives limited insight into efficacy, particularly when lack of efficacy is then interpreted as implying the absence of an underlying effect. Thought still needs to go into alternative ways of analysing the outcome of interventions, particularly within consultations, for MUS which reflect the non-independence and likely non-normal distribution of outcome events.

Self similarity across severities and processes

A fundamental characteristic of complex systems is that of self similarity at different scales. This may be in different time scales – for instance the variability in physiological data is similar over widely differing timescales(198) or over different levels, for instance the continuity of the distribution of consultation patterns between infrequent and very frequent consulters(14).

In the case of MUS two forms of self-similarity are particularly relevant. The first is the similarity of illness experience and consultation behaviour across different symptoms, independently of whether they are of organic or functional origin. In this form of self-similarity, the association of physical symptoms and psychological distress holds regardless of cause(70). The second form of self similarity involves the ways patients organise information about their illness and experience, with prototypes(296) and clusters of events and meanings rather than the strict timeline approach used in medical history taking. Exploring the patient’s story as clusters of

events which are similar in outline but different in detail may be an appropriate alternative to constraining a complicated experience within a simplistic disease label or time-line structure of cause and effect.

Both forms of similarity pose problems for research. Firstly it is clear that setting inclusion and exclusion criteria becomes difficult in the presence of large overlaps between processes. If similar symptoms and levels of severity can emerge from a variety of underlying physiological and cognitive processes, careful specification of inclusion criteria of only one kind (e.g. symptom counts or syndrome labels) will lead to overestimations of homogeneity. Secondly, if the presentation of mixed emotional and physical symptoms and meaning is the norm for patients(293), selecting a particular clinical style for one group of problems (with GPs suddenly switching to a specific reattribution mode) becomes artificial. It may be that GPs need to adopt a different overall style which is more open to cues of emotional distress and which, by letting the patient tell their story in a less constricted way, picks up on ideas about possible causes. While potentially effective(306), such a model probably comes at the price of requiring extra time and a loss of efficiency from the perspective of general practice with priorities of identifying and investigating organic pathology, chronic disease management and lifestyle risk factor modification ahead of helping people deal with difficult lives and bodies.

Conclusion

In this chapter I have argued that reframing MUS as complex adaptive disorders, with maladaptive physiological and psychological processes rather than psychopathology is justified from both theory and evidence. The effects of this in practice are still to be tested, however the model converges with empirical evidence in terms of the importance of enabling explanations(40) and the danger of oversimplifying the situation(33;305). The challenge may be to find explanations for patients that embrace complexity while permitting them to find coherent meaning and achieve therapeutic action.

Chapter 23 Conclusions

The title of this thesis, “complex illness – variation and causality in medically unexplained symptoms” raises three challenges. To categorise medically unexplained symptoms in a way which makes sense for patients and doctors in primary care; to collect and interpret the complicated data from patients’ own experience; and to formulate these as a complex illness.

Medically unexplained symptoms

By reviewing the literature I have shown that the concept of somatisation is inadequate to explain symptoms which represent the experience of physiological processes in the context of a personal emotional and cognitive biography. In addition, it appears that the processes of making sense of symptoms and seeking help for them are broadly similar to those used with “explained symptoms”.

Nonetheless it is clear that some patients suffering functional symptom disorders have high levels of emotional distress which often reaches psychiatric caseness. Despite this, many seek non-psychiatric meanings for their symptoms, sometimes using problem solving strategies which have served them well in the past, or avoiding particularly threatening biographical territory.

From interview data, symptom specific labels were shown both to be helpful (at least in part) to those who used them, but also weak in that patterns of variation, association and interpretation of symptoms did not make sense in terms of the label used.

Quantitative assessment of diaries of self reported symptom severity showed that mind and body interact at a measurable level in patients with MUS, but that uncovering this requires sophisticated analysis of detailed personal data rather than reliance on self-perception or easily refutable argument by analogy.

Variation and causality

Through an original analysis of a complex multivariate set of self-report time series data I have demonstrated considerable day to day variation in the severity of functional physical symptoms and psychological states in a heterogeneous group of patients with MUS.

Correlations between symptoms and psychological variables differ both between symptom–variable pairs and between individuals recording the same symptom–variable pairs. The concurrent associations of anxiety and stress with a wide range of MUS were small, this was in contrast with low mood which was more strongly associated with symptoms.

A novel analysis of sequential processes, using granger causality for the first time with self report data, shows evidence for bidirectional influence between symptoms and psychological states, with unexpected but plausible findings that some functional symptoms, particularly visceral pain, engender more emotional distress than others.

The relative weakness of the associations with anxiety and stress are important given the conventional notion of MUS and somatisation as stress related illnesses.

Complex illness

In this thesis I have dealt with the notion of complex illness on two levels. Firstly it is clear that the interactions between mind and body, or cognitions and physiology, are complicated and multi-directional. Reducing functional disorders to misattribution of emotional difficulty is shown to be an over-simplification. Participants' narratives confirmed the notion of chaotic illness and the search for meaning experienced by patients trying to make sense of their illness and also highlighted the inadequate explanations and support provided by health professionals.

Secondly, the notion of complex illness was explored from the perspective of the science of complex systems. Participants' own data was shown to possess lower complexity (entropy) than predicted from surrogate control data and consultation patterns within healthcare systems and the interactions of physiological and cognitive processes within individuals support a complex systems approach.

In considering complex illnesses, as researchers and clinicians, we need to rethink the management of unexplained symptoms: we must recognise the day to day complexity, unpredictability and uncertainty inherent in these disorders; we may have to accept that conventional “black-box” interventions limited to patients with MUS will be of limited value in primary care and find ways of integrating emotional and physical components of symptoms and distress across the symptom spectrum;

and we must live with the knowledge that illness trajectories are long, and that small changes at one stage will usually achieve little, but will sometimes have sustained and surprising consequences.

Complex illness – variation and causality in medically unexplained symptoms

When we label the experience of patients with a range of illnesses, due more to dysfunction than pathology, with classifications developed for diseases with more simple cause, we create problems for patients and doctors. This is particularly troublesome when, as this research suggests, one of the commonly invoked pathogenic processes (“stress”) is so weakly correlated with the experience of symptoms. The research in this thesis shows a marked variation in symptoms which cannot be accounted for by simple models of cause and effect, and provides evidence from the experience of psychological states and physical symptoms for a “two way street” between mind and body which complements that demonstrated by physiological research.

In proposing a solution for the problems of simplification, I have argued that a critical realist, complex systems, approach offers new ways of conceptualising medically unexplained symptoms as dynamic adaptations to threats. This reformulation casts new light on the difficulties of researching these problems, and offers opportunities to help patients make sense of, and ultimately deal with, their illness.

The next stage of this research is to develop interventions which acknowledge the complexity of these illnesses, both in their content and the way in which they are evaluated. For now, Engel’s principle of enabling patients to understand their body and be understood as individuals(304) seems as good a guide to clinical practice as any.

Reference List

- (1) Royal College of General Practitioners. The future general practitioner: learning and teaching. London: RCGP, 1976.
- (2) National Health Service E. NHS (General Medical Services) Regulations 2004. London: The Stationery Office, 2004.
- (3) Leventhal H, Meyer D, Nerenz D. The common sense representation of illness danger. In: Rachman S, editor. Contributions to medical psychology (volume II). New York: Pergamon Press, 1980: 7-30.
- (4) Damasio A. The feeling of what happens; body, emotion and the making of consciousness. London: Vintage, 2000.
- (5) Helman CG. Culture, health and illness. 4th ed. Oxford: Butterworth-Heinemann, 2000.
- (6) Horne R, James D, Petrie K, Weinman J, Vincent R. Patients' interpretation of symptoms as a cause of delay in reaching hospital during acute myocardial infarction. *Heart* 2000; 83(4):388-393.
- (7) Zola IK. Pathways to the doctor-from person to patient. *Soc Sci Med* 1973; 7(9):677-689.
- (8) Campbell SM, Roland MO. Why do people consult the doctor? *Family Practice* 1996; 13:75-83.
- (9) Verbrugge LM. Triggers of symptoms and health care. *Soc Sci Med* 1985; 20(9):855-876.
- (10) Hannay DR. The symptom iceberg: a study of community health. London: Routledge & Kegan Paul, 1979.
- (11) Jordan K, Ong BN, Croft P. Previous consultation and self reported health status as predictors of future demand for primary care. *J Epidemiol Community Health* 2003; 57(2):109-113.
- (12) Little P, Somerville J, Williamson I, Warner G, Moore M, Wiles R et al. Psychosocial, lifestyle, and health status variables in predicting high attendance among adults. *British Journal of General Practice* 2001; 51(473):987-994.
- (13) Pollock K, Grime J, Mechanic D. Patients' perceptions of entitlement to time in general practice consultations for depression: qualitative study * Commentary: Managing time appropriately in primary care. *BMJ* 2002; 325(7366):687.
- (14) Love T, Burton C. General practice as a complex system: a novel analysis of consultation data. *Fam Pract* 2005; 22(3):347-352.
- (15) Little P, Gould C, Williamson I, Warner G, Gantley M, Kinmonth AL. Reattendance and complications in a randomised trial of prescribing strategies for

- sore throat: the medicalising effect of prescribing antibiotics. *BMJ* 1997; 315(7104):350-352.
- (16) Robinson A, Lee V, Kennedy A, Middleton L, Rogers A, Thompson DG et al. A randomised controlled trial of self-help interventions in patients with a primary care diagnosis of irritable bowel syndrome. *Gut* 2006; 55(5):643-648.
 - (17) Peveler R, Kilkenny L, Kinmonth AL. Medically unexplained physical symptoms in primary care: a comparison of self-report screening questionnaires and clinical opinion. *Journal of Psychosomatic Research* 1997; 42(3):245-252.
 - (18) Kroenke K. Patients presenting with somatic complaints: epidemiology, psychiatric comorbidity and management. *Int J Methods Psychiatr Res* 2003; 12(1):34-43.
 - (19) Nimnuan C, Hotopf M, Wessely S. Medically unexplained symptoms: an epidemiological study in seven specialities. *J Psychosom Res* 2001; 51(1):361-367.
 - (20) Wessely S, Nimnuan M, Sharpe M. Functional Somatic Syndromes: one or many. *Lancet* 1999; 354:936-939.
 - (21) Sharpe M, Carson A. "Unexplained" somatic symptoms, functional syndromes, and somatization: do we need a paradigm shift? [see comments.]. [Review] [34 refs]. *Annals of Internal Medicine* 2001; 134(9 Pt 2):926-930.
 - (22) Stone J, Wojcik W, Durrance D, Carson A, Lewis S, MacKenzie L et al. What should we say to patients with symptoms unexplained by disease? The "number needed to offend". *BMJ* 2002; 325(7378):1449-1450.
 - (23) Escobar JI, Gara M, Silver RC, Waitzkin H, Holman A, Compton W. Somatisation disorder in primary care. *British Journal of Psychiatry* 1998; 173:262-266.
 - (24) Barsky AJ, Wyshak G, Klerman GL, Latham KS. The prevalence of hypochondriasis in medical outpatients. *Social Psychiatry & Psychiatric Epidemiology* 1990; 25(2):89-94.
 - (25) Butler CC, Evans M, Greaves D, Simpson S. Medically unexplained symptoms: the biopsychosocial model found wanting. *J R Soc Med* 2004; 97(5):219-222.
 - (26) Lillrank A. Back pain and the resolution of diagnostic uncertainty in illness narratives. *Soc Sci Med* 2003; 57(6):1045-1054.
 - (27) Werner A, Isaksen LW, Malterud K. 'I am not the kind of woman who complains of everything': illness stories on self and shame in women with chronic pain. *Soc Sci Med* 2004; 59(5):1035-1045.
 - (28) Cartwright N. *The dappled world, a study of the boundaries of science*. Cambridge: Cambridge University Press, 1999.
 - (29) Helman CG. Cultural aspects of time and ageing. Time is not the same in every culture and every circumstance; our views of aging also differ. *EMBO Rep* 2005; 6 Spec No:S54-S58.

- (30) Wileman L, May C, Chew-Graham CA. Medically unexplained symptoms and the problem of power in the primary care consultation: a qualitative study. *Fam Pract* 2002; 19(2):178-182.
- (31) Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. *Ann Behav Med* 2002; 24(3):236-243.
- (32) Bradburn NM, Rips LJ, Shevell SK. Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 1987; 236(4798):157-161.
- (33) Nettleton S, Watt I, O'Malley L, Duffey P. Understanding the narratives of people who live with medically unexplained illness. *Patient Education & Counseling* 2005; . 56(2).
- (34) Mabeck CE, Olesen F. Metaphorically transmitted diseases. How do patients embody medical explanations? *Family Practice* 1997; 14(4):271-278.
- (35) Marcum JA. Biomechanical and phenomenological models of the body, the meaning of illness and quality of care. *Med Health Care Philos* 2004; 7(3):311-320.
- (36) Lucock MP, Morley S, White C, Peake MD. Responses of consecutive patients to reassurance after gastroscopy: results of self administered questionnaire survey. *BMJ* 1997; 315(7108):572-575.
- (37) Howard L, Wessely S, Leese M, Page L, McCrone P, Husain K et al. Are investigations anxiolytic or anxiogenic? A randomised controlled trial of neuroimaging to provide reassurance in chronic daily headache. *J Neurol Neurosurg Psychiatry* 2005; 76(11):1558-1564.
- (38) Spiegel BM, Gralnek IM, Bolus R, Chang L, Dulai GS, Naliboff B et al. Is a negative colonoscopy associated with reassurance or improved health-related quality of life in irritable bowel syndrome? *Gastrointest Endosc* 2005; 62(6):892-899.
- (39) Donovan JL, Blake RD. Qualitative study of interpretation of reassurance among patients attending rheumatology clinics: "just a touch of arthritis, doctor?". *BMJ* 2001; 320:541-544.
- (40) Salmon P, Peters S, Stanley I. Patients' perceptions of medical explanations for somatisation disorders: qualitative analysis. *BMJ* 1999; 318:372-376.
- (41) Salmon P, Humphris GM, Ring A, Davies JC, Dowrick CF. Why do primary care physicians propose medical care to patients with medically unexplained symptoms? A new method of sequence analysis to test theories of patient pressure. *Psychosom Med* 2006; 68(4):570-577.
- (42) Dowrick CF, Ring A, Humphris GM, Salmon P. Normalisation of unexplained symptoms by general practitioners: a functional typology. *Br J Gen Pract* 2004; 54(500):165-170.

- (43) Goldberg DP, Bridges K. Somatic presentations of psychiatric illness in primary care setting. [Review] [15 refs]. *Journal of Psychosomatic Research* 1988; 32(2):137-144.
- (44) Dowrick C, Buchan I. Twelve month outcome of depression in general practice: does detection or disclosure make a difference? *BMJ* 1995; 311(7015):1274-1276.
- (45) Rief W, Sharpe M. Somatoform disorders-new approaches to classification, conceptualization, and treatment. *J Psychosom Res* 2004; 56(4):387-390.
- (46) Fink P, Rosendal M, Toft T. Assessment and treatment of functional disorders in general practice: the extended reattribution and management model--an advanced educational program for nonpsychiatric doctors. *Psychosomatics* 2002; 43(2):93-131.
- (47) Bak P. *How Nature Works: The Science of Self-Organised Criticality*. London: Oxford University Press, 1996.
- (48) Burton C. Introduction to complexity. In: Sweeney K, Griffiths F, editors. *Complexity and Healthcare*. Oxford: Radcliffe, 2002: 1-18.
- (49) Dancy CP, Taghavi M, Fox RJ. The relationship between daily stress and symptoms of irritable bowel: a time-series approach. *J Psychosom Res* 1998; 44(5):537-545.
- (50) Helman CG. Limits of biomedical explanation. *Lancet* 1991; 337(8749):1080-1083.
- (51) Woloshynowych M, Valori P, Salmon P. General practice patients' beliefs about their symptoms. *British Journal of General Practice* 1998; 48:885-889.
- (52) Kroenke K, Swindle R. Cognitive-behavioral therapy for somatization and symptom syndromes: a critical review of controlled clinical trials. [Review] [57 refs]. *Psychotherapy & Psychosomatics* 2000; 69(4):205-215.
- (53) Walker EA, Katon WJ, Jemelka RP, Roy-Bryne PP. Comorbidity of gastrointestinal complaints, depression, and anxiety in the Epidemiologic Catchment Area (ECA) Study. *Am J Med* 1992; 92(1A):26S-30S.
- (54) Lipowski ZJ. Somatization: the concept and its clinical application. *American Journal of Psychiatry* 1988; 145:1358-1368.
- (55) American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC, American Psychiatric Association. Fourth Edition ed. Washington DC: American Psychiatric Association, 1994.
- (56) Bridges K, Goldberg D. Somatic presentation of depressive illness in primary care. *Journal of the Royal College of General Practitioners - Occasional Paper* 1992;(36):9-11.
- (57) Katon WJ, Walker EA. Medically unexplained symptoms in primary care. [Review] [54 refs]. *Journal of Clinical Psychiatry* 1998; 59 Suppl 20:15-21.

- (58) Peters S, Stanley I, Rose M, Salmon P. Patients with medically unexplained symptoms: sources of patients authority and implications for demands on medical care. *Social Science & Medicine* 2001; 46:559-565.
- (59) Kroenke K. Physical symptom disorder: a simpler diagnostic category for somatization-spectrum conditions. *J Psychosom Res* 2006; 60(4):335-339.
- (60) de Lange FP, Kalkman JS, Bleijenberg G, Hagoort P, van der Meer JW, Toni I. Gray matter volume reduction in the chronic fatigue syndrome. *Neuroimage* 2005; 26(3):777-781.
- (61) Cleare AJ. The HPA axis and the genesis of chronic fatigue syndrome. *Trends Endocrinol Metab* 2004; 15(2):55-59.
- (62) Hansson E. Could chronic pain and spread of pain sensation be induced and maintained by glial activation? *Acta Physiol (Oxf)* 2006; 187(1-2):321-327.
- (63) Bennett R. Fibromyalgia: present to future. *Curr Rheumatol Rep* 2005; 7(5):371-376.
- (64) Mumford DB, Devereux TA, Maddy PJ, Johnston JV. Factors leading to the reporting of 'functional' somatic symptoms by general practice attenders. *British Journal of General Practice* 1991; 41(352):454-458.
- (65) Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. *Fam Pract* 2006; 23(4):414-420.
- (66) Escobar JI, Waitzkin H, Silver RC, Gara M, Holman A. Abridged somatization: a study in primary care. *Psychosomatic Medicine* 1998; 60(4):466-472.
- (67) Kirmayer LJ, Robbins JM. Three forms of somatization in primary care: prevalence, co-occurrence, and sociodemographic characteristics. *J Nerv Ment Dis* 1991; 179(11):647-655.
- (68) Kisely S, Goldberg D, Simon G. A comparison between somatic symptoms with and without clear organic cause: results of an international study. *Psychological Medicine* 1997; 27(5):1011-1019.
- (69) Fink P, Sorensen L, Engberg M, Holm M, Munk-Jorgensen P. Somatization in primary care. Prevalence, health care utilization, and general practitioner recognition. *Psychosomatics* 1999; 40(4):330-338.
- (70) Kroenke K, Spitzer RL, Williams JB, Linzer M, Hahn SR, deGruy FV, III et al. Physical symptoms in primary care. Predictors of psychiatric disorders and functional impairment. *Archives of Family Medicine* 1994; 3(9):774-779.
- (71) Melville DI. Descriptive clinical research and medically unexplained physical symptoms. *Journal of Psychosomatic Research* 1987; 31:359-365.
- (72) Palsson N. Functional somatic symptoms and hypochondriasis among general practice patients: a pilot study. *Acta Psychiatrica Scandinavica* 1988; 78(2):191-197.

- (73) Pilowsky I, Smith QP, Katsikitis M. Illness behaviour and general practice utilisation: a prospective study. *Journal of Psychosomatic Research* 1987; 31(2):177-183.
- (74) Scicchitano J, Lovell P, Pearce R, Marley J, Pilowsky I. Illness behavior and somatization in general practice. *Journal of Psychosomatic Research* 1996; 41(3):247-254.
- (75) Posse M, Hallstrom T. Depressive disorders among somatizing patients in primary health care. *Acta Psychiatrica Scandinavica* 1998; 98(3):187-192.
- (76) Toft T, Fink P, Oernboel E, Christensen K, Frostholt L, Olesen F. Mental disorders in primary care: prevalence and co-morbidity among disorders. results from the functional illness in primary care (FIP) study. *Psychol Med* 2005; 35(8):1175-1184.
- (77) de Waal MW, Arnold IA, Eekhof JA, Van Hemert AM. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. *Br J Psychiatry* 2004; 184:470-476.
- (78) Aggarwal VR, McBeth J, Zakrzewska JM, Lunt M, Macfarlane GJ. The epidemiology of chronic syndromes that are frequently unexplained: do they have common associated factors? *Int J Epidemiol* 2006; 35(2):468-476.
- (79) Lobo A, Garcia-Campayo J, Campos R, Marcos G, Perez-Echeverria MJ. Somatisation in primary care in Spain: I. Estimates of prevalence and clinical characteristics. Working Group for the Study of the Psychiatric and Psychosomatic Morbidity in Zaragoza. *British Journal of Psychiatry* 1996; 168(3):344-348.
- (80) Kirmayer LJ, Robbins JM, Dworkind M, Yaffe MJ. Somatization and the recognition of depression and anxiety in primary care. *American Journal of Psychiatry* 1993; 150(5):734-741.
- (81) Munk-Jorgensen P, Fink P, Brevik JI, Dalgard OS, Engberg M, Hansson L et al. Psychiatric morbidity in primary public health care: a multicentre investigation. Part II. Hidden morbidity and choice of treatment. *Acta Psychiatrica Scandinavica* 1997; 95(1):6-12.
- (82) Feder A, Olfson M, Gameroff M, Fuentes M, Shea S, Lantigua RA et al. Medically unexplained symptoms in an urban general medicine practice. *Psychosomatics* 2001; 42(3):261-268.
- (83) Gara MA, Silver RC, Escobar JI, Holman A, Waitzkin H. A hierarchical classes analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms. *Psychiatry Research* 1998; 81(1):77-86.
- (84) Robbins JM, Kirmayer LJ, Hemami S. Latent variable models of functional somatic distress. *Journal of Nervous & Mental Disease* 1997; 185(10):606-615.
- (85) Liu G, Clark MR, Eaton WW. Structural factor analyses for medically unexplained somatic symptoms of somatization disorder in the Epidemiologic Catchment Area study. *Psychol Med* 1997; 27(3):617-626.

- (86) Deary IJ. A taxonomy of medically unexplained symptoms. *Journal of Psychosomatic Research* 1999; 47(1):51-59.
- (87) Craig TK, Drake H, Mills K, Boardman AP. The South London Somatisation Study. II. Influence of stressful life events, and secondary gain. *British Journal of Psychiatry* 1994; 165(2):248-258.
- (88) Hotopf M, Wadsworth M, Wessely S. Is "somatisation" a defense against the acknowledgment of psychiatric disorder? *Journal of Psychosomatic Research* 2001; 50(3):119-124.
- (89) Cameron L, Leventhal EA, Leventhal H. Seeking medical care in response to symptoms and life stress [see comments]. *Psychosomatic Medicine* 1995; 57(1):37-47.
- (90) Fry RP, Crisp AH, Beard RW. Sociopsychological factors in chronic pelvic pain: a review. [Review] [53 refs]. *Journal of Psychosomatic Research* 1997; 42(1):1-15.
- (91) Hotopf M, Mayou R, Wadsworth M, Wessely S. Temporal relationships between physical symptoms and psychiatric disorder. Results from a national birth cohort. *British Journal of Psychiatry* 1998; 173:255-261.
- (92) Garcia-Campayo J, Sanz-Carrillo C. A review of the differences between somatizing and psychologizing patients in primary care. *International Journal of Psychiatry in Medicine* 1999; 29(3):337-345.
- (93) Kirmayer LJ, Robbins JM. Patients who somatize in primary care: a longitudinal study of cognitive and social characteristics. *Psychological Medicine* 1996; 26(5):937-951.
- (94) Cape J, McCulloch Y. Patients' reasons for not presenting emotional problems in general practice consultations. *British Journal of General Practice* 1999; 49(448):875-879.
- (95) Bower P, West R, Tylee A, Hann M. Symptom attribution and the recognition of psychiatric morbidity. *Journal of Psychosomatic Research* 2000; 48(2):157-160.
- (96) Jyvasjarvi S, Joukamaa M, Vaisanen E, Larivaara P, Kivela SL, Keinanen-Kiukaanniemi S. Alexithymia, hypochondriacal beliefs, and psychological distress among frequent attenders in primary health care. *Comprehensive Psychiatry* 1999; 40(4):292-298.
- (97) Noyes R, Jr., Langbehn DR, Happel RL, Sieren LR, Muller BA. Health Attitude Survey. A scale for assessing somatizing patients. *Psychosomatics* 1999; 40(6):470-478.
- (98) MacLeod AK, Haynes C, Sensky T. Attributions about common bodily sensations: their associations with hypochondriasis and anxiety. - *Psychol Med* 1998 Jan;28(1):225-8 1998;(1):225-228.
- (99) Robbins JM, Kirmayer LJ. Transient and persistent hypochondriacal worry in primary care. *Psychological Medicine* 1996; 26(3):575-589.

- (100) Speckens AE, Spinhoven P, Sloekers PP, Bolk JH, Van Hemert AM. A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. *Journal of Psychosomatic Research* 1996; 40(1):95-104.
- (101) Speckens AE, Van Hemert AM, Spinhoven P, Bolk JH. The diagnostic and prognostic significance of the Whitely Index, the Illness Attitude Scales and the Somatosensory Amplification Scale. *Psychological Medicine* 1996; 26(5):1085-1090.
- (102) Barsky AJ, Ahern DK, Bailey ED, Delamater BA. Predictors of persistent palpitations and continued medical utilization. *J Fam Pract* 1996; 42(5):465-472.
- (103) Barsky AJ, Barnett MC, Cleary PD. Hypochondriasis and panic disorder. Boundary and overlap. *Archives of General Psychiatry* 1994; 51(11):918-925.
- (104) Sensky T, MacLeod AK, Rigby MF. Causal attributions about common somatic sensations among frequent general practice attenders. *Psychological Medicine* 1996; 26(3):641-646.
- (105) Coia P, Morley S. Medical reassurance and patients' responses. *Journal of Psychosomatic Research* 1998; 45:377-386.
- (106) Williams S, Weinman J, Dale J, Newman S. Patient expectations: what do primary care patients want from the GP and how far does meeting expectations affect patient satisfaction? *Fam Pract* 1995; 12(2):193-201.
- (107) Salmon P, Woloshynowych M, Valori R. The measurement of beliefs about physical symptoms in English general practice patients. *Social Science & Medicine* 1996; 42:1561-1567.
- (108) Simon GE, Gureje O. Stability of somatization disorder and somatization symptoms among primary care patients. *Archives of General Psychiatry* 1999; 56(1):90-95.
- (109) Garcia-Campayo J, Larrubia J, Lobo A, Perez-Echeverria MJ, Campos R. Attribution in somatizers: stability and relationship to outcome at 1-year follow-up. Grupo Morbilidad Psiquica y Psicomatica de Zaragoza (GMPPZ). *Acta Psychiatrica Scandinavica* 1997; 95(5):433-438.
- (110) van Dulmen AM, Fennis JF, Mookink HG, Bleijenberg G. The relationship between complaint-related cognitions in referred patients with irritable bowel syndrome and subsequent health care seeking behaviour in primary care. *Family Practice* 1996; 13(1):12-17.
- (111) Crombez G, Eccleston C, Van den BA, Goubert L, Van Houdenhove B. Hypervigilance to pain in fibromyalgia: the mediating role of pain intensity and catastrophic thinking about pain. *Clin J Pain* 2004; 20(2):98-102.
- (112) Lackner JM, Quigley BM, Blanchard EB. Depression and abdominal pain in IBS patients: the mediating role of catastrophizing. *Psychosom Med* 2004; 66(3):435-441.

- (113) Gracely RH, Geisser ME, Giesecke T, Grant MA, Petzke F, Williams DA et al. Pain catastrophizing and neural responses to pain among persons with fibromyalgia. *Brain* 2004; 127(Pt 4):835-843.
- (114) Petrie K, Moss-Morris R, Weinman J. The impact of catastrophic beliefs on functioning in chronic fatigue syndrome. *J Psychosom Res* 1995; 39(1):31-37.
- (115) Turner JA, Mancl L, Aaron LA. Pain-related catastrophizing: a daily process study. *Pain* 2004; 110(1-2):103-111.
- (116) Smeets RJ, Vlaeyen JW, Kester AD, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *J Pain* 2006; 7(4):261-271.
- (117) Potts SG, Bass CM. Psychological morbidity in patients with chest pain and normal or near-normal coronary arteries: a long-term follow-up study. *Psychological Medicine* 1995; 25(2):339-347.
- (118) Neal RD, Heywood PL, Morley S. 'I always seem to be there'--a qualitative study of frequent attenders. *Br J Gen Pract* 2000; 50(458):716-723.
- (119) Morriss RK, Gask L, Ronalds C, Downes-Grainger E, Thompson H, Goldberg D. Clinical and patient satisfaction outcomes of a new treatment for somatized mental disorder taught to general practitioners. *British Journal of General Practice* 1999; 49(441):263-267.
- (120) Morriss R, Gask L, Ronalds C, Downes-Grainger E, Thompson H, Leese B et al. Cost-effectiveness of a new treatment for somatized mental disorder taught to GPs. *Family Practice* 1998; 15(2):119-125.
- (121) Downes-Grainger E, Morriss R, Gask L, Faragher B. Clinical factors associated with short-term changes in outcome of patients with somatized mental disorder in primary care. *Psychological Medicine* 1998; 28(3):703-711.
- (122) Lidbeck J. Group therapy for somatization disorders in general practice: effectiveness of a short cognitive-behavioural treatment model. *Acta Psychiatrica Scandinavica* 1997; 96(1):14-24.
- (123) McLeod CC, Budd MA, McClelland DC. Treatment of somatization in primary care. *General Hospital Psychiatry* 1997; 19(4):251-258.
- (124) Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain* 1999; 80(1-2):1-13.
- (125) Sharpe M. Cognitive behavior therapy for functional somatic complaints. The example of chronic fatigue syndrome. *Psychosomatics* 1997; 38(4):356-362.
- (126) Hadhazy VA, Ezzo J, Creamer P, Berman BM. Mind-body therapies for the treatment of fibromyalgia. A systematic review. *Journal of Rheumatology* 2000; 27(12):2911-2918.

- (127) Speckens AE, Van Hemert AM, Spinhoven P, Hawton KE, Bolk JH, Rooijmans HG. Cognitive Behavioural therapy for medically unexplained symptoms: a randomised controlled trial. *BMJ* 1995; 311:1328-1332.
- (128) Guthrie E, Creed F, Dawson D, Tomenson B. A randomised controlled trial of psychotherapy in patients with refractory irritable bowel syndrome. *British Journal of Psychiatry* 1993; 163:315-321.
- (129) Schilte AF, Portegijs PJM, Blankenstein AH, Van der Horst HE, Latour MBF, van Eijk JT et al. Randomised controlled trial of disclosure of emotionally important events in somatisation in primary care. *British Medical Journal* 2001; 323(7304):86-89.
- (130) O'Malley PG, Jackson JL, Santoro J, Tomkins G, Balden E, Kroenke K. Antidepressant therapy for unexplained symptoms and symptom syndromes. *J Fam Pract* 1999; 48(12):980-990.
- (131) Goldberg D, Gask L, O'Dowd T. The treatment of somatization: teaching techniques of reattribution. *Journal of Psychosomatic Research* 1989; 33(6):689-695.
- (132) Blankenstein AH, Van der Horst HE, Schilte AF, de Vries D, Zaat JO, Andre KJ et al. Development and feasibility of a modified reattribution model for somatising patients, applied by their own general practitioners. *Patient Educ Couns* 2002; 47(3):229-235.
- (133) Larisch A, Schweickhardt A, Wirsching M, Fritzsche K. Psychosocial interventions for somatizing patients by the general practitioner: a randomized controlled trial. *J Psychosom Res* 2004; 57(6):507-514.
- (134) Schweickhardt A, Larisch A, Fritzsche K, Schweickhardt A, Larisch A, Fritzsche K. Differentiation of somatizing patients in primary care: why the effects of treatment are always moderate. *Journal of Nervous & Mental Disease* 2005; 193(12):813-819.
- (135) Rosendal M, Bro F, Sokolowski I, Fink P, Toft T, Olesen F. A randomised controlled trial of brief training in assessment and treatment of somatisation: effects on GPs' attitudes. *Fam Pract* 2005; 22(4):419-427.
- (136) Morriss R, Dowrick C, Salmon P, Peters S, Rogers A, Dunn G et al. Turning theory into practice: rationale, feasibility and external validity of an exploratory randomized controlled trial of training family practitioners in reattribution to manage patients with medically unexplained symptoms (the MUST). *Gen Hosp Psychiatry* 2006; 28(4):343-351.
- (137) van der Feltz-Cornelis CM, van Oppen P, Ader HJ, van Dyck R. Randomised controlled trial of a collaborative care model with psychiatric consultation for persistent medically unexplained symptoms in general practice. *Psychother Psychosom* 2006; 75(5):282-289.
- (138) Little P, Everitt H, Williamson I, Warner G, Moore M, Gould C et al. Observational study of effect of patient centredness and positive approach on outcomes of general practice consultations. *BMJ* 2001; 323(7318):908-911.

- (139) Lunghi ME, Miller PM, McQuillan WM. Psycho-social factors in osteoarthritis of the hip. *J Psychosom Res* 1978; 22(1):57-63.
- (140) Persson LO, Sjoberg L. Mood and somatic symptoms. *J Psychosom Res* 1987; 31(4):499-511.
- (141) Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. *Annu Rev Psychol* 2003; 54:579-616.
- (142) Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. *BMJ* 2002; 324(7347):1193-1194.
- (143) Conner T. Experience Sampling Resource Page. <http://psychiatry.uhc.edu/people/postdocs/conner/esm.php> . 27-10-2005. 27-7-2006.
- (144) Palmblad M, Tiplady B. Electronic diaries and questionnaires: designing user interfaces that are easy for all patients to use. *Qual Life Res* 2004; 13(7):1199-1207.
- (145) Schwartz JE, Stone AA. Strategies for analyzing ecological momentary assessment data. *Health Psychol* 1998; 17(1):6-16.
- (146) Shiffman S, Gwaltney CJ, Balabanis MH, Liu KS, Paty JA, Kassel JD et al. Immediate antecedents of cigarette smoking: an analysis from ecological momentary assessment. *J Abnorm Psychol* 2002; 111(4):531-545.
- (147) Affleck G, Zautra A, Tennen H, Armeli S. Multilevel daily process designs for consulting and clinical psychology: a preface for the perplexed. *J Consult Clin Psychol* 1999; 67(5):746-754.
- (148) Feiler S, Muller KG, Muller A, Dahlhaus R, Eich W. Using interaction graphs for analysing the therapy process. *Psychother Psychosom* 2005; 74(2):93-99.
- (149) Brown KW, Moskowitz DS. It's a function of time: a review of the process approach to behavioral medicine research. *Ann Behav Med* 1998; 20(2):109-117.
- (150) Gendreau M, Hufford MR, Stone AA. Measuring clinical pain in chronic widespread pain: selected methodological issues. *Best Pract Res Clin Rheumatol* 2003; 17(4):575-592.
- (151) Koop A, Mosges R. The use of handheld computers in clinical trials. *Control Clin Trials* 2002; 23(5):469-480.
- (152) Giffin NJ, Ruggiero L, Lipton RB, Silberstein SD, Tvedskov JF, Olesen J et al. Premonitory symptoms in migraine: an electronic diary study. *Neurology* 2003; 60(6):935-940.
- (153) Honkoop PC, Sorbi MJ, Godaert GL, Spierings EL. High-density assessment of the IHS classification criteria for migraine without aura: a prospective study. *Cephalalgia* 1999; 19(4):201-206.

- (154) Saito M, Kumano H, Yoshiuchi K, Kokubo N, Ohashi K, Yamamoto Y et al. Symptom profile of multiple chemical sensitivity in actual life. *Psychosom Med* 2005; 67(2):318-325.
- (155) Liszka-Hackzell JJ, Martin DP. An analysis of the relationship between activity and pain in chronic and acute low back pain. *Anesth Analg* 2004; 99(2):477-81, table.
- (156) Affleck G, Urrows S, Tennen H, Higgins P, Abeles M. Sequential daily relations of sleep, pain intensity, and attention to pain among women with fibromyalgia. *Pain* 1996; 68(2-3):363-368.
- (157) Litcher-Kelly L, Stone AA, Broderick JE, Schwartz JE. Associations among pain intensity, sensory characteristics, affective qualities, and activity limitations in patients with chronic pain: a momentary, within-person perspective. *J Pain* 2004; 5(8):433-439.
- (158) Aaron LA, Turner JA, Mancl L, Brister H, Sawchuk CN. Electronic diary assessment of pain-related variables: is reactivity a problem? *J Pain* 2005; 6(2):107-115.
- (159) Turner JA, Mancl L, Aaron LA. Brief cognitive-behavioral therapy for temporomandibular disorder pain: Effects on daily electronic outcome and process measures. *Pain* 2005; 117(3):377-387.
- (160) Aaron LA, Mancl L, Turner JA, Sawchuk CN, Klein KM. Reasons for missing interviews in the daily electronic assessment of pain, mood, and stress. *Pain* 2004; 109(3):389-398.
- (161) Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. *Pain* 2003; 104(1-2):343-351.
- (162) Stone AA, Broderick JE, Shiffman SS, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain* 2004; 107(1-2):61-69.
- (163) Litt MD, Shafer D, Napolitano C. Momentary mood and coping processes in TMD pain. *Health Psychol* 2004; 23(4):354-362.
- (164) Roelofs J, Peters ML, Patijn J, Schouten EG, Vlaeyen JW. Electronic diary assessment of pain-related fear, attention to pain, and pain intensity in chronic low back pain patients. *Pain* 2004; 112(3):335-342.
- (165) Affleck G, Tennen H, Zautra A, Urrows S, Abeles M, Karoly P. Women's pursuit of personal goals in daily life with fibromyalgia: a value-expectancy analysis. *J Consult Clin Psychol* 2001; 69(4):587-596.
- (166) Zautra A, Smith B, Affleck G, Tennen H. Examinations of chronic pain and affect relationships: applications of a dynamic model of affect. *J Consult Clin Psychol* 2001; 69(5):786-795.

- (167) Peters ML, Sorbi MJ, Kruse DA, Kerssens JJ, Verhaak PF, Bensing JM. Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. *Pain* 2000; 84(2-3):181-192.
- (168) Viane I, Crombez G, Eccleston C, Devulder J, De Corte W. Acceptance of the unpleasant reality of chronic pain: effects upon attention to pain and engagement with daily activities. *Pain* 2004; 112(3):282-288.
- (169) Affleck G, Apter A, Tennen H, Reisine S, Barrows E, Willard A et al. Mood states associated with transitory changes in asthma symptoms and peak expiratory flow. *Psychosom Med* 2000; 62(1):61-68.
- (170) Kop WJ, Lyden A, Berlin AA, Ambrose K, Olsen C, Gracely RH et al. Ambulatory monitoring of physical activity and symptoms in fibromyalgia and chronic fatigue syndrome. *Arthritis Rheum* 2005; 52(1):296-303.
- (171) Gaertner J, Elsner F, Pollmann-Dahmen K, Radbruch L, Sabatowski R. Electronic pain diary: a randomized crossover study. *J Pain Symptom Manage* 2004; 28(3):259-267.
- (172) Williams DA, Gendreau M, Hufford MR, Groner K, Gracely RH, Clauw DJ. Pain assessment in patients with fibromyalgia syndrome: a consideration of methods for clinical trials. *Clin J Pain* 2004; 20(5):348-356.
- (173) Lauritsen K, Degl' IA, Hendel L, Praest J, Lytje MF, Clemmensen-Rotne K et al. Symptom recording in a randomised clinical trial: paper diaries vs. electronic or telephone data capture. *Control Clin Trials* 2004; 25(6):585-597.
- (174) Kreindler D, Levitt A, Woolridge N, Lumsden CJ. Portable mood mapping: the validity and reliability of analog scale displays for mood assessment via hand-held computer. *Psychiatry Res* 2003; 120(2):165-177.
- (175) Quinn P, Goka J, Richardson H. Assessment of an electronic daily diary in patients with overactive bladder. *BJU Int* 2003; 91(7):647-652.
- (176) Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials* 2003; 24(2):182-199.
- (177) Jamison RN, Gracely RH, Raymond SA, Levine JG, Marino B, Herrmann TJ et al. Comparative study of electronic vs. paper VAS ratings: a randomized, crossover trial using healthy volunteers. *Pain* 2002; 99(1-2):341-347.
- (178) Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain* 2001; 91(3):277-285.
- (179) Johannes CB, Crawford SL, Woods J, Goldstein RB, Tran D, Mehrotra S et al. An electronic menstrual cycle calendar: comparison of data quality with a paper version. *Menopause* 2000; 7(3):200-208.
- (180) Lewis B, Lewis D, Cumming G. Frequent measurement of chronic pain: an electronic diary and empirical findings. *Pain* 1995; 60(3):341-347.

- (181) Vendrig AA, Lousberg R. Within-person relationships among pain intensity, mood and physical activity in chronic pain: a naturalistic approach. *Pain* 1997; 73(1):71-76.
- (182) Price JR, Blake F. Chronic pelvic pain: the assessment as therapy. *J Psychosom Res* 1999; 46(1):7-14.
- (183) Chatfield C. *The analysis of time series - an introduction*. 6th ed. Boca Raton: CRC Press, 2004.
- (184) Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 1969; 37:424-438.
- (185) Kaminski M, Ding M, Truccolo WA, Bressler SL. Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern* 2001; 85(2):145-157.
- (186) Blinowska KJ, Kus R, Kaminski M. Granger causality and information flow in multivariate processes. *Phys Rev E Stat Nonlin Soft Matter Phys* 2004; 70(5 Pt 1):050902.
- (187) Schreiber T. Measuring information transfer. *Phys Rev Lett* 2000; 85(2):461-464.
- (188) Field AP. Meta-analysis of correlations. <http://www.sussex.ac.uk/Users/andyf/meta.pdf>. 2006. 3-5-0006.
- (189) Engel GL. The need for a new medical model: a challenge for biomedicine. *Science* 1977; 196(4286):129-136.
- (190) Cohen J, Stewart I. *The Collapse of Chaos: discovering simplicity in a complex world*. London: Penguin, 1994.
- (191) Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci U S A* 2002; 99 Suppl 3:7280-7287.
- (192) Plsek PE, Greenhalgh T. Complexity science: The challenge of complexity in health care. *BMJ* 2001; 323(7313):625-628.
- (193) Mauroy B, Filoche M, Weibel ER, Sapoval B. An optimal bronchial tree may be dangerous. *Nature* 2004; 427(6975):633-636.
- (194) Pincus SM. Assessing serial irregularity and its implications for health. *Ann N Y Acad Sci* 2001; 954:245-267.
- (195) Goldberger AL, Amaral LAN, Hausdorff JM, Ivanov PC, Peng CK, Stanley HE. Fractal dynamics in physiology: Alterations with disease and aging. *PNAS* 2002; 99(90001):2466-2472.
- (196) Watters PA, Martin F. A method for estimating long-range power law correlations from the electroencephalogram. *Biol Psychol* 2004; 66(1):79-89.
- (197) Smethurst DP, Williams HC. Power laws. Are hospital waiting lists self-regulating? *Nature* 2001; 410(6829):652-653.

- (198) Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of biological signals. *Phys Rev E Stat Nonlin Soft Matter Phys* 2005; 71(2 Pt 1):021906.
- (199) Stam CJ. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin Neurophysiol* 2005; 116(10):2266-2301.
- (200) Miller WL, Crabtree BF, McDaniel R, Stange KC. Understanding change in primary care practice using complexity theory. *J Fam Pract* 1998; 46(5):369-376.
- (201) Weisstein EW. "Chaos." From MathWorld--A Wolfram Web Resource. <http://mathworld.wolfram.com/Chaos.html> . 2006. 26-7-0006.

Ref Type: Electronic Citation

- (202) Pincus S. Approximate entropy (ApEn) as a complexity measure. *Chaos* 1995; 5(1):110-117.
- (203) Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 2000; 278(6):H2039-H2049.
- (204) Goldberger AL, Peng CK, Lipsitz LA. What is physiologic complexity and how does it change with aging and disease? *Neurobiol Aging* 2002; 23(1):23-26.
- (205) Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett* 2002; 89(6):068102.
- (206) Goldberger AL. Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *Lancet* 1996; 347(9011):1312-1314.
- (207) Gottschalk A, Bauer MS, Whybrow PC. Evidence of chaotic mood variation in bipolar disorder. *Arch Gen Psychiatry* 1995; 52(11):947-959.
- (208) Abasolo D, Hornero R, Espino P, Alvarez D, Poza J. Entropy analysis of the EEG background activity in Alzheimer's disease patients. *Physiol Meas* 2006; 27(3):241-253.
- (209) Bahrami B, Seyedsadjadi R, Babadi B, Noroozian M. Brain complexity increases in mania. *Neuroreport* 2005; 16(2):187-191.
- (210) Woyshtville MJ, Lackamp JM, Eisengart JA, Gilliland JA. On the meaning and measurement of affective instability: clues from chaos theory. *Biol Psychiatry* 1999; 45(3):261-269.
- (211) Yeragani VK, Pohl R, Mallavarapu M, Balon R. Approximate entropy of symptoms of mood: an effective technique to quantify regularity of mood. *Bipolar Disord* 2003; 5(4):279-286.
- (212) Heiby EM, Pagano IS, Blaine DD, Nelson K, Heath RA. Modeling Unipolar Depression as a Chaotic Process. *Psychological Assessment* 2003; . 15(3).
- (213) Thayer JF, Friedman BH. Stop that! Inhibition, sensitization, and their neurovisceral concomitants. *Scand J Psychol* 2002; 43(2):123-130.

- (214) Nimnuan C, Rabe-Hesketh S, Wessely S, Hotopf M. How many functional somatic syndromes? *J Psychosom Res* 2001; 51(4):549-557.
- (215) Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990; 33(2):160-172.
- (216) Snaith RP, Zigmond AS. The hospital anxiety and depression scale. *Br Med J (Clin Res Ed)* 1986; 292(6516):344.
- (217) Barsky AJ, Wyshak G. Hypochondriasis and somatosensory amplification. *British Journal of Psychiatry* 1990; 157:404-409.
- (218) Fink P, Ewald H, Jensen J, Sorensen L, Engberg M, Holm M et al. Screening for somatization and hypochondriasis in primary care and neurological in-patients: a seven-item scale for hypochondriasis and somatization. *Journal of Psychosomatic Research* 1999; 46(3):261-273.
- (219) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2006.
- (220) The Python Programming Language. 2006.
- (221) MSBVAR: Bayesian Vector Autoregression Models for R. 2006.
- (222) Hunter JE, Schmidt FL. *Methods of Meta-Analysis: correcting error and bias in research findings*. 2nd edition ed. Thousand Oaks: Sage Publications, 2004.
- (223) Psychometric Package - Applied Psychometric Theory for R. 2006.
- (224) Whitener EM. Trust in the face of conflict: the role of managerial trustworthy behavior and organizational context. *J Appl Psychol* 1990; 75(3):315-321.
- (225) Schreiber T, Schmitz A. Surrogate time series. *Physica D* 2000; 142:346.
- (226) Hegger R, Kantz H, Schreiber T. Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos* 1999; 9:413.
- (227) Buck R, Morley S. A daily process design study of attentional pain control strategies in the self-management of cancer pain. *Eur J Pain* 2006; 10(5):385-398.
- (228) Parsons HM. What happened at Hawthorne? *Science* 1974; 183(4128):922-932.
- (229) Undeland M, Malterud K. Diagnostic work in general practice: more than naming a disease. *Scand J Prim Health Care* 2002; 20(3):145-150.
- (230) Ong BN, Hooper H. Comparing clinical and lay accounts of the diagnosis and treatment of back pain. *Sociol Health Illn* 2006; 28(2):203-222.
- (231) Nettleton S. 'I just want permission to be ill': Towards a sociology of medically unexplained symptoms. *Social Science & Medicine* 2006; . 62(5).

- (232) Kleinman A. Patients and healers in the context of culture. London: University of California Press, 1980.
- (233) Frostholm L, Fink P, Oernboel E, Christensen KS, Toft T, Olesen F et al. The uncertain consultation and patient satisfaction: the impact of patients' illness perceptions and a randomized controlled trial on the training of physicians' communication skills. *Psychosom Med* 2005; 67(6):897-905.
- (234) Madden S, Sim J. Creating meaning in fibromyalgia syndrome. *Soc Sci Med* 2006; 63(11):2962-2973.
- (235) Spence M, Moss-Morris R, Chalder T. The Behavioural Responses to Illness Questionnaire (BRIQ): a new predictive measure of medically unexplained symptoms following acute infection. *Psychol Med* 2005; 35(4):583-593.
- (236) Kirmayer LJ, Robbins JM, Paris J. Somatoform disorders: personality and the social matrix of somatic distress. *J Abnorm Psychol* 1994; 103(1):125-136.
- (237) Barry CA, Stevenson FA, Britten N, Barber N, Bradley CP. Giving voice to the lifeworld. More humane, more effective medical care? A qualitative study of doctor-patient communication in general practice. *Soc Sci Med* 2001; 53(4):487-505.
- (238) Dancey CP, Whitehouse A, Painter J, Backhouse S. The relationship between hassles, uplifts and irritable bowel syndrome: a preliminary study. *J Psychosom Res* 1995; 39(7):827-832.
- (239) Suls J, Wan CK, Blanchard EB. A multilevel data-analytic approach for evaluation of relationships between daily life stressors and symptomatology: patients with irritable bowel syndrome. *Health Psychol* 1994; 13(2):103-113.
- (240) McGorm K. Patients frequently referred from primary care to hospital outpatients for medically unexplained symptoms. University of Edinburgh, 2006.
- (241) O'Dowd TC. Five years of heartsink patients in general practice. *BMJ* 1988; 297(6647):528-530.
- (242) Salmon P, Dowrick CF, Ring A, Humphris GM. Voiced but unheard agendas: qualitative analysis of the psychosocial cues that patients with unexplained symptoms present to general practitioners. *Br J Gen Pract* 2004; 54(500):171-176.
- (243) Rief W, Martin A, Rauh E, Zech T, Bender A. Evaluation of general practitioners' training: how to manage patients with unexplained physical symptoms. *Psychosomatics* 2006; 47(4):304-311.
- (244) Sharpe M, Mayou R, Walker J. Bodily symptoms: new approaches to classification. *J Psychosom Res* 2006; 60(4):353-356.
- (245) McKenzie N, Marks I. Quick rating of depressed mood in patients with anxiety disorders. *Br J Psychiatry* 1999; 174:266-269.
- (246) Rosendal M, Bro F, Fink P, Christensen KS, Olesen F, Rosendal M et al. Diagnosis of somatisation: effect of an educational intervention in a cluster

randomised controlled trial.[see comment]. *British Journal of General Practice* 2003; 53(497):917-922.

- (247) Lake DE, Richman JS, Griffin MP, Moorman JR. Sample entropy analysis of neonatal heart rate variability. *Am J Physiol Regul Integr Comp Physiol* 2002; 283(3):R789-R797.
- (248) Olsson I, Mykletun A, Dahl AA. The Hospital Anxiety and Depression Rating Scale: a cross-sectional study of psychometrics and case finding abilities in general practice. *BMC Psychiatry* 2005; 5:46.
- (249) Banks SM, Kerns RD. Explaining high rates of depression in chronic pain: a diathesis-stress framework. *Psychological Bulletin* 1996; 119(1):95-110.
- (250) Peveler R, Katona C, Wessely S, Dowrick C. Painful symptoms in depression: under-recognised and under-treated? *Br J Psychiatry* 2006; 188:202-203.
- (251) Schoenen J. Depression in tension-type headache sufferers: bystander or villain? *Pain* 2004; 111(3):225.
- (252) Carroll LJ, Cassidy JD, Cote P. Depression as a risk factor for onset of an episode of troublesome neck and low back pain. *Pain* 2004; 107(1-2):134-139.
- (253) Epstein SA, Kay G, Clauw D, Heaton R, Klein D, Krupp L et al. Psychiatric disorders in patients with fibromyalgia. A multicenter investigation. *Psychosomatics* 1999; 40(1):57-63.
- (254) Wessely S, Chalder T, Hirsch S, Wallace P, Wright D. Psychological symptoms, somatic symptoms, and psychiatric disorder in chronic fatigue and chronic fatigue syndrome: a prospective study in the primary care setting. *Am J Psychiatry* 1996; 153(8):1050-1059.
- (255) Henningsen P, Zimmermann T, Sattel H. Medically unexplained physical symptoms, anxiety, and depression: a meta-analytic review. *Psychosom Med* 2003; 65(4):528-533.
- (256) Brown KW, Moskowitz DS. Does unhappiness make you sick? The role of affect and neuroticism in the experience of common physical symptoms. *J Pers Soc Psychol* 1997; 72(4):907-917.
- (257) Crane C, Martin M, Johnston D, Goodwin GM. Does depression influence symptom severity in irritable bowel syndrome? Case study of a patient with irritable bowel syndrome and bipolar disorder. *Psychosom Med* 2003; 65(5):919-923.
- (258) Tennen H, Affleck G, Zautra A. Depression history and coping with chronic pain: a daily process analysis. *Health Psychol* 2006; 25(3):370-379.
- (259) Conner TS, Tennen H, Zautra AJ, Affleck G, Armeli S, Fifield J. Coping with rheumatoid arthritis pain in daily life: Within-person analyses reveal hidden vulnerability for the formerly depressed. *Pain* 2006.

- (260) Harris RE, Williams DA, McLean SA, Sen A, Hufford M, Gendreau RM et al. Characterization and consequences of pain variability in individuals with fibromyalgia. *Arthritis Rheum* 2005; 52(11):3670-3674.
- (261) Janke EA, Holroyd KA, Romanek K. Depression increases onset of tension-type headache following laboratory stress. *Pain* 2004; 111(3):230-238.
- (262) Craig TK, Boardman AP, Mills K, Daly-Jones O, Drake H. The South London Somatisation Study. I: Longitudinal course and the influence of early life experiences [see comments]. *British Journal of Psychiatry* 1993; 163:579-588.
- (263) Schilte AF, Portegijs PJ, Blankenstein AH, Latour MB, van Eijk JT, Knottnerus JA. Indicators of childhood adversity in somatisation in general practice. *Scand J Prim Health Care* 2001; 19(4):232-236.
- (264) Hotopf M, Mayou R, Wadsworth M, Wessely S. Childhood risk factors for adults with medically unexplained symptoms: results from a national birth cohort study. *American Journal of Psychiatry* 1999; 156(11):1796-1800.
- (265) Ciechanowski PS, Walker EA, Katon WJ, Russo JE. Attachment theory: a model for health care utilization and somatization. *Psychosom Med* 2002; 64(4):660-667.
- (266) Hotopf M. Preventing somatization. *Psychol Med* 2004; 34(2):195-198.
- (267) Skapinakis P, Lewis G, Mavreas V. Temporal relations between unexplained fatigue and depression: longitudinal data from an international study in primary care. *Psychosom Med* 2004; 66(3):330-335.
- (268) Kupers R, Faymonville ME, Laureys S. The cognitive modulation of pain: hypnosis- and placebo-induced analgesia. *Prog Brain Res* 2005; 150:251-269.
- (269) Eisenberger NI, Lieberman MD. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends Cogn Sci* 2004; 8(7):294-300.
- (270) Giesecke T, Gracely RH, Williams DA, Geisser ME, Petzke FW, Clauw DJ. The relationship between depression, clinical pain, and experimental pain in a chronic pain cohort. *Arthritis Rheum* 2005; 52(5):1577-1584.
- (271) Chang L, Berman S, Mayer EA, Suyenobu B, Derbyshire S, Naliboff B et al. Brain responses to visceral and somatic stimuli in patients with irritable bowel syndrome with and without fibromyalgia. *Am J Gastroenterol* 2003; 98(6):1354-1361.
- (272) Sarchielli P, Alberti A, Candelieri A, Floridi A, Capocchi G, Calabresi P. Glial cell line-derived neurotrophic factor and somatostatin levels in cerebrospinal fluid of patients affected by chronic migraine and fibromyalgia. *Cephalalgia* 2006; 26(4):409-415.
- (273) Vlaeyen JW, Morley S. Active despite pain: the putative role of stop-rules and current mood. *Pain* 2004; 110(3):512-516.
- (274) de Gier M, Peters ML, Vlaeyen JW. Fear of pain, physical performance, and attentional processes in patients with fibromyalgia. *Pain* 2003; 104(1-2):121-130.

- (275) McCracken LM, Eccleston C. A prospective study of acceptance of pain and patient functioning with chronic pain. *Pain* 2005; 118(1-2):164-169.
- (276) Ring A, Dowrick C, Humphris G, Salmon P. Do patients with unexplained physical symptoms pressurise general practitioners for somatic treatment? A qualitative study. *BMJ* 2004; 328(7447):1057.
- (277) Craig AD. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci* 2002; 3(8):655-666.
- (278) Jarrett ME, Burr RL, Cain KC, Hertig V, Weisman P, Heitkemper MM. Anxiety and depression are related to autonomic nervous system function in women with irritable bowel syndrome. *Dig Dis Sci* 2003; 48(2):386-394.
- (279) Kapur N, Hunt I, Lunt M, McBeth J, Creed F, Macfarlane G. Psychosocial and illness related predictors of consultation rates in primary care--a cohort study. *Psychol Med* 2004; 34(4):719-728.
- (280) Halder SL, McBeth J, Silman AJ, Thompson DG, Macfarlane GJ. Psychosocial risk factors for the onset of abdominal pain. Results from a large prospective population-based study. *Int J Epidemiol* 2002; 31(6):1219-1225.
- (281) Thakkar RR, McCanne TR. The effects of daily stressors on physical health in women with and without a childhood history of sexual abuse. *Child Abuse Negl* 2000; 24(2):209-221.
- (282) Thomas E, Moss-Morris R, Faquhar C. Coping with emotions and abuse history in women with chronic pelvic pain. *J Psychosom Res* 2006; 60(1):109-112.
- (283) Craig TK, Bialas I, Hodson S, Cox AD. Intergenerational transmission of somatization behaviour: 2. Observations of joint attention and bids for attention. *Psychol Med* 2004; 34(2):199-209.
- (284) DeLongis A, Folkman S, Lazarus RS. The impact of daily stress on health and mood: psychological and social resources as mediators. *J Pers Soc Psychol* 1988; 54(3):486-495.
- (285) Holtzman S, Newth S, DeLongis A. The role of social support in coping with daily pain among patients with rheumatoid arthritis. *J Health Psychol* 2004; 9(5):677-695.
- (286) Pini S, Piccinelli M, Zimmermann-Tansella C. Social problems as factors affecting medical consultation: a comparison between general practice attenders and community probands with emotional distress. *Psychol Med* 1995; 25(1):33-41.
- (287) Traue HC, Kosarz P. Everyday stress and Crohn's disease activity: a time series analysis of 20 single cases. *Int J Behav Med* 1999; 6(2):101-119.
- (288) Epstein RM, Shields CG, Meldrum SC, Fiscella K, Carroll J, Carney PA et al. Physicians' responses to patients' medically unexplained symptoms. *Psychosom Med* 2006; 68(2):269-276.

- (289) Stevens JA, Wan CK, Blanchard EB. The short-term natural history of irritable bowel syndrome: a time-series analysis. *Behav Res Ther* 1997; 35(4):319-326.
- (290) Graugaard PK, Eide H, Finset A. Interaction analysis of physician-patient communication: the influence of trait anxiety on communication and outcome. *Patient Educ Couns* 2003; 49(2):149-156.
- (291) Finset A, Graugaard PK, Holgersen K. Salivary cortisol response after a medical interview: the impact of physician communication behaviour, depressed affect and alexithymia. *Patient Educ Couns* 2006; 60(2):115-124.
- (292) Miceli M, Castelfranchi C. The plausibility of defensive projection. *Journal of the Theory of Social Behaviour* 2003; 33:279-301.
- (293) Ring A, Dowrick CF, Humphris GM, Davies J, Salmon P. The somatising effect of clinical consultation: what patients and doctors say and do not say when patients present medically unexplained physical symptoms. *Soc Sci Med* 2005; 61(7):1505-1515.
- (294) Rief W, Nanke A, Emmerich J, Bender A, Zech T. Causal illness attributions in somatoform disorders: associations with comorbidity and illness behavior. *J Psychosom Res* 2004; 57(4):367-371.
- (295) Stern L, Kirmayer LJ. Knowledge structures in illness narratives: development and reliability of a coding scheme. *Transcult Psychiatry* 2004; 41(1):130-142.
- (296) Dominice Dao M, Kirmayer LJ, Groleau D, Young A. Mingling mind and body: patients' complex networks of meaning in the absence of medical diagnosis. *J Psychosom Res* 2006; 61(3):410.
- (297) Young A. When rational men fall sick: an inquiry into some assumptions made by medical anthropologists. *Cult Med Psychiatry* 1981; 5(4):317-335.
- (298) Kleinman A. On illness meanings and clinical interpretation: not 'rational man', but a rational approach to man the sufferer/man the healer. *Cult Med Psychiatry* 1981; 5(4):373-377.
- (299) Wikipedia, . Critical Realism. wikipedia.org . 2006. 15-10-2006.

Ref Type: Electronic Citation

- (300) Pawson R, Tilley N. *Realistic Evaluation*. London: Sage, 1997.
- (301) Ciechanowski P, Sullivan M, Jensen M, Romano J, Summers H. The relationship of attachment style to depression, catastrophizing and health care utilization in patients with chronic pain. *Pain* 2003; 104(3):627-637.
- (302) Rhodes LA, McPhillips-Tangum CA, Markham C, Klenk R. The power of the visible: the meaning of diagnostic tests in chronic back pain. *Soc Sci Med* 1999; 48(9):1189-1203.
- (303) Klapow JC, Schmidt SM, Taylor LA, Roller P, Li Q, Calhoun JW et al. Symptom management in older primary care patients: feasibility of an experimental, written self-disclosure protocol. *Annals of Internal Medicine* 2001; 134(9 Pt 2):905-911.

- (304) Engel GL. How much longer must medicine's science be bound by a seventeenth century world view? In: White KL, editor. *The task of medicine, dialogue at Wickenburg*. Menlo Park, CA: Henry J. Kaiser Family Foundation, 1988: 113-136.
- (305) Kirmayer LJ. Improvisation and authority in illness meaning. *Cult Med Psychiatry* 1994; 18(2):183-214.
- (306) Matalon A, Nahmani T, Rabin S, Maoz B, Hart J, Matalon A et al. A short-term intervention in a multidisciplinary referral clinic for primary care frequent attenders: description of the model, patient characteristics and their use of medical resources. *Family Practice* 2002; 19(3):251-256.
- (307) Salmon P, Ring A, Dowrick CF, Humphris GM. What do general practice patients want when they present medically unexplained symptoms, and why do their doctors feel pressurized? *J Psychosom Res* 2005; 59(4):255-260.
- (308) Luyten P, Blatt SJ, Van Houdenhove B, Corveleyn J. Depression research and treatment: Are we skating to where the puck is going to be? *Clin Psychol Rev* 2006.
- (309) Surtees PG, Wainwright NW, Khaw KT. Resilience, misfortune, and mortality: evidence that sense of coherence is a marker of social stress adaptive capacity. *J Psychosom Res* 2006; 61(2):221-227.
- (310) Soderlund A, Malterud K. Why did I get chronic fatigue syndrome? A qualitative interview study of causal attributions in women patients. *Scand J Prim Health Care* 2005; 23(4):242-247.
- (311) de Lange FP, Kalkman JS, Bleijenberg G, Hagoort P, van der Werf SP, van der Meer JW et al. Neural correlates of the chronic fatigue syndrome--an fMRI study. *Brain* 2004; 127(Pt 9):1948-1957.
- (312) Dantzer R, Kelley KW. Twenty years of research on cytokine-induced sickness behavior. *Brain Behav Immun* 2006.
- (313) Lyall M, Peakman M, Wessely S. A systematic review and critical evaluation of the immunology of chronic fatigue syndrome. *J Psychosom Res* 2003; 55(2):79-90.

Appendix A STUDY PROTOCOL

Functional Bodily Symptoms Diary Study *Subject Identification Code.....*

A. Pre-recruitment: participating GPs

GP monitors consultations and incoming mail and notes names of patients presenting with functional symptoms.

GP reviews names and checks records for entry criteria (FBSD 2a)

GP writes to patient, enclosing study leaflet and inviting to take part in study.

Stages B to D take place consecutively in one to one interviews between (potential) subjects and researcher on GP's premises.

B. Recruitment

Introduction and verbal explanation of aims and rationale of study

Potential subject handed further copy of study leaflet, researcher ensures that potential subject has read it and invites questions. Palm diary demonstration can be given at this stage if potential subject wishes.

Researcher goes through entry checklist with potential subject to confirm eligibility. (FBSD 2a)

Consent discussed and, if appropriate, signed (FBSD 1)

Subject study number allocated

C. Baseline measures

Patient choice of symptoms for monitoring and categorisation of symptom syndromes using specific criteria obtained through brief clinical history (FBSD 6). There will be no physical examination or blood tests carried out.

Symptom Questionnaire: Somatic Symptom Inventory (FBSD 3)

Illness concern Questionnaire: Whiteley 7 (FBSD 5)

General Mental Health Questionnaire: Hospital Anxiety & Depression Scale (FBSD 4)

D. Configuration and explanation of computer diary

Subject is handed a demonstration diary to try out (no data from this will be recorded, lockouts to prevent duplicate entry will be disabled) until they feel comfortable handling the diary.

Researcher configures symptom diary with subject's study identifier and chosen symptoms

Subject is loaned, handheld computer diary and arrangements made for first review and for telephone support of the diary.

Subject given paper event sheet to record significant events (FSBD 7)

Stage E recurs at 3-4 week intervals, in one to one interviews between researcher and subjects

E. Review of progress and collection of data

Welcome and brief discussion of wellbeing

Discussion of any difficulty either with the meaning of questions or the use of the diary

Electronic retrieval of data from the subject's diary

Collection of paper event sheet, provision of new sheet

Arrangements made for next review

Stage F occurs after the final data collection

F. Final interview

1. After final stage E (approximately three months after starting) research GP collects computer diary and paper diaries from patient

2. Long (approximately one hour) consultation in which research GP and patient jointly explore patient's symptoms and results of electronic diary. Outcome of this will be notified to patient and patient's GP in writing.

Appendix B CONSENT FORM

SANQUHAR HEALTH CENTRE

Dr Chris Burton

FBSD1
24 February 2004

Patient Identification Code

CONSENT FORM

Title of Project: Functional Bodily Symptom Diary Study 2004-6

Name of Researcher: Dr Chris Burton

Please initial each box

1. I confirm that I have read and understand the information sheet dated
for the above study and have had the opportunity to ask questions.
2. I understand that my participation is voluntary and that I am free to withdraw at any time,
without giving any reason, without my medical care or legal rights being affected.
3. I understand that the data collected by me will be analysed and published for
research purposes but in such a form as to make my identity unknown.
4. I agree to take part in the above study.
5. I understand that the handheld computer diary supplied to me remains the property of
Sanquhar Health Centre. I will take good care of it and return it at the end of the study.

Name of Participant

Date

Signature

Name of Person taking consent
(if different from researcher)

Date

Signature

Researcher

Date

Signature

February 2004

1 for participant; 1 for researcher; 1 to be kept with practice notes

Phone (01659) 50221 Station Road Sanquhar Dumfriesshire DG4 6BT Fax (01659) 58116
E-mail chrisburton@medicine21.com

Appendix C HAD SCALE

Functional Bodily Symptoms Diary Study HAD Scale

FBSD 4
22 November 2002

Subject identification number.....

For each item select one option which comes closest to the way you have been feeling in the last week

I feel tense or 'wound up'
Most of the time.....
A lot of the time
Time to time, occasionally
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I feel as if I am slowed down
Nearly all the time
Very often
Sometimes
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I still enjoy the things I used to enjoy
Definitely as much
Not quite so much
Only a little
Hardly at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I get a sort of frightened feeling like "butterflies"
in the stomach
Not at all
Occasionally
Quite often
Very often

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I get a sort of frightened feeling as if something
awful is about to happen
Very definitely and quite badly
Yes, but not too badly
A little, but it doesn't worry me
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I have lost interest in my appearance
Definitely
I don't take as much care as I should
A little, but it doesn't worry me
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I can laugh and see the funny side of things
As much as I always could
Not quite so much now
Definitely not so much now
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I feel restless as if I have to be on the move
Very much indeed
Quite a lot
Not very much
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

Worrying thoughts go through my mind
A great deal of the time
A lot of the time
From time to time but not too often
Only occasionally

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I look forward with enjoyment to things
As much as ever I did
Rather less than I used to
Definitely less than I used to
Hardly at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I feel cheerful
Not at all
Not often
Sometimes
Most of the time

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I get sudden feelings of panic
Very often indeed
Quite often
Not very often
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I can sit at ease and feel relaxed
Definitely
Usually
Not often
Not at all

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

I can enjoy a good book or TV programme
Often
Sometimes
Not often
Very seldom

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

Appendix E SOMATIC SYMPTOM INVENTORY

Please indicate which of these symptoms you have **ever** had which were **present for more than a few weeks** (either continuously or on and off) and **bad enough** that you had to **see a doctor, or take medicine, or change what you do or eat**

Vomiting (other than during pregnancy)	Yes	No
Abdominal pain (except period pain in women)	Yes	No
Nausea (other than travel sickness)	Yes	No
Abdominal bloating and wind	Yes	No
Diarrhoea	Yes	No
Burning pain in rectum (back passage)	Yes	No
Intolerance of several different foods	Yes	No
Chest pain	Yes	No
Breathlessness when not exerting yourself	Yes	No
Palpitations	Yes	No
Back pain	Yes	No
Pain in arms and legs	Yes	No
Joint pain	Yes	No
Pain during urination	Yes	No
Headaches	Yes	No
Other bodily pains not listed above	Yes	No
Forgetfulness	Yes	No
Double vision	Yes	No
Blurred vision	Yes	No
Blindness	Yes	No
Fainting or loss of consciousness	Yes	No
Seizure or convulsion	Yes	No
Trouble walking	Yes	No
Difficulty swallowing	Yes	No
Loss of voice	Yes	No
Deafness	Yes	No
Paralysis or muscle weakness	Yes	No
Difficulty passing urine	Yes	No
Pain during intercourse	Yes	No
Burning sensation in sexual organs except during intercourse	Yes	No
No interest in sex	Yes	No
Impotence	Yes	No
Painful periods	Yes	No
Irregular periods	Yes	No
Excessive bleeding with periods	Yes	No

Appendix F SELF REFERRAL GP CHECKLIST

Inclusion / Exclusion Criteria

Dear Doctor

re

Patient Name

Patient Address

This person has contacted us as a volunteer participant in this study. He / she has given written consent to us to discuss his / her symptoms, initially by telephone, and has agreed that we check his/her suitability for this study with you, their named GP.

Contra-indications are as follows

The patient's predominant physical symptoms are due to a physical illness in which physical tests have shown unequivocal abnormalities (e.g. radiology, blood tests)

The patient has, or has in the past had, major physical disease such as coronary heart disease or cancer

The patient has had significant depression, or treatment for depression in the last 6 months. Minor depression without early wakening or anhedonia etc, or use of antidepressants for pain or other symptoms are not absolute contraindications.

The patient has ongoing drug or psychological treatment for other major psychiatric illness and

In addition if you believe that anything in the patient's history suggests a risk to personal safety from a researcher visiting them at home please notify me.

Unless you notify us of any contra-indication to his / her participation in the study within the next two weeks we will arrange to meet them for a recruitment interview.

You can notify me either by letter, fax (01659 58116) or email (christopher.burton@nhs.net)

If you wish to discuss the case with me personally please feel free to phone on 01659 50221.

Appendix G Instructions for handheld computer diary.

Thank you for agreeing to take part in this study using a handheld computer symptom diary. This information is to remind you how to use the diary. The system has been designed to be straightforward to use and hopefully you should have no difficulty. If you have a problem please phone Sanquhar Health Centre (01659) 50221 and leave a message for the research team to call you back about your handheld computer diary.

Taking care of the computer

Your handheld personal computer (HPC) is provided on loan for the duration of the project (usually around three months) thereafter you must return it to the research team. Please take care with the HPC which should be handled in the same way as a portable CD player or electronic game.

Switching on

The simplest way to switch on your HPC is to press the right hand metallic button which has a little picture of a notepad on it. This switches on the power and loads the VASco programme.



Press this button to start

Using your HPC

Handheld computers work by touching the screen with a blunt pointer called a stylus. The stylus for your HPC is in a groove on the right hand side of the machine and should be kept there whenever you are not using it. Please always use the stylus rather than a pen or pencil to operate the HPC. If you lose it please let us know and we will provide a replacement.

To touch an area of the screen a light single press will usually suffice – imagine dotting an “i” rather than drawing a line or making a mark. If you do this on an active area of the screen you will promptly either get a response (clicking an “OK” area will make the screen change in some way) or see a mark.

Writing on HPCs is a little more difficult than simply marking a point so the VASco programme has been designed to avoid this or any other complicated use of the stylus.

Reminders

When you receive your HPC it will be pre-set to sound an alarm on two occasions each day. These will be chosen by you for convenience. If an alarm goes off you will see a screen asking you what to do. Touch the “OK” area of the screen to switch off or, if you wish to be reminded in a few minutes touch the snooze area of the screen

Although it is good if you can enter data at the same times each day, you can enter data on the HPC at other times. To stop you entering data twice at the same time, however, there is a timer which stops the programme if it is less than four hours since your last entry.

If you have been unable to enter data for a few days, when you switch on you will get a reminder that this has been the case. Whilst there are bound to be times and days when you cannot enter data, the more data we have from you, the more use it is for analysis later.

Starting VASco

Regardless of whether your HPC is already on, for instance if it has switched on for an alarm, or off, press the right hand metallic button which has a little picture of a notepad on it. This will start the VASco programme and you will see the welcome screen.

If you get a reminder that it is several days since you last entered data then touch the “OK” area with the stylus.

If you get a warning that it is less than four hours since you last entered data you do not need to do anything more as the HPC will switch itself off.

Entering data in VASco

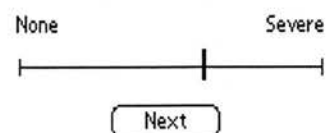
The main VASco data entry is in eight questions. These are about three physical symptoms you agreed at the start of the study, your energy level, anxiety, depression, how much stress you are under and how concerned or worried you are about your symptoms. You should answer these on how you have been feeling **over the last few hours**.

For each question you should mark a point on the line indicating your answer. For instance if you had selected back pain as one of your symptoms you would see the screen in the picture below

Back Pain

How much have you been bothered by back pain?

Please mark a point on the line between severe back pain and no back pain at all.



To enter your reading, use the stylus to mark a point along the line. If you just touch the line in one spot, the programme will display a line like the one in the picture.

If you want to change it simply touch a spot in a different place on the line. When you are happy with your answer touch the “Next” area. This will lead you to the next screen.

The questions in VASco are the same every time you run it, but each time the order in which they occur is shuffled. That way you never know which question is coming next. Sometimes you will find that at the end you are asked one of the questions again. This does not mean there is a problem, it is simply there as part of a checking mechanism in the programme.

If you accidentally touch the “Next” area before you have entered a point on the line, VASco will stop and ask you what you want to do. If you meant to answer that question tap the “No” area and you will get another chance. If you really don’t want to answer that question on the particular day tap the “Yes” area.

Joint Pain

How much have you been bothered by joint pain?

Please mark a point on the line between severe joint pain and no

Warning



You have not marked the line. Skip this and go to the next page?

No

Yes

Concern about symptoms

How concerned are you about your symptoms?

Please mark a point on the scale between very concerned and not concerned at all.

Not at all Very



Next

One of the eight questions VASco asks is about how concerned you are about your symptoms. Concern, in this context, is not a measure of how bad the symptoms are, but of how you feel about them. For instance if you had a cough for a couple of days which you knew was part of a simple cold you may have been up all night with it but not be particularly concerned. On the other hand if you had the same cough but had spent a fortnight with pneumonia last year you may be rather more concerned.

Once the symptom questions are finished, you will be asked a simple set of questions about contact with doctors or nurses. If the answer to any of these simply tap with the stylus in the box and a tick will appear. To get rid of it simply tap again.

If there is an important health or personal event it would be helpful if you could make a short note about it in a pen and paper notebook.

When you have finished

After the last question, VASco will tell you your session is finished and you can put the HPC down. It will switch itself off after a few seconds..... until the next reminder!

If you have any questions about these instructions please let us know.

Events

Since you last entered data have you....

- Seen a doctor or nurse?
- Spoken to a doctor or nurse?
- Made an appointment?

- Had an important health or personal event (diary)?

Next

Appendix H Meta-analysis Computer Script (python)

```
'''Poolcorrs 4: a modified version of poolcorrs
collects correlations from file,
groups them by variable pair
runs Hunter & Schmidt Random effects Meta-analysis
with optional weighting by sample size and test-retest reliability (ICC)
'''
```

```
from string import split
from Numeric import array,arrayrange, zeros
from Numeric import shape
from Numeric import Float
from Numeric import log
from Numeric import exp
from time import gmtime, strftime,sleep
```

```
CorrGrid = zeros((26,152),Float)
```

```
def loadSlists(fname,path):
```

```
'''
creates list of each subject ID, sample size, and ICC
'''
listID = []
listN=[]
listE=[]
```

```
f=open(path + fname)
for line in f:
a = split(line)
listID.append(int(a[0]))
listN.append(int(a[1]))
listE.append(float(a[2]))
return listID,listN,listE
```

```
def loadXlists():
```

```
'''
Creates list of all possible interactions between named variables, with
each pairing appearing only once.
'''
```

```
listL =['Fatigue','Stress','Depression','Anxiety','Concern','Muscle Pain',
'Headache', 'Joint Pain', 'Bowel Trouble', 'Abdominal Pain', 'Indigestion',
'Nausea', 'Throat', 'Chest Pain', 'Pelvic Pain', 'Weakness', 'Back Pain',
'Numbness', 'Palpitations']
list2=[]
for i in range(12):
for j in range(len(listL)):
if i<j:
list2.append(listL[i][:4]+'/'+listL[j][:4])
#print len(list2)
return list2
```

```
def matchXname(listX,name):
```

```
# finds the position in listX of the pairing supplied as name.
# the pairing can be in either order: Aaaa/Bbbb or Bbbb/Aaaa
# returns the position in the list
```

```

k = name
k2 = k[5:]+ '/' + k[:4]
try: #see if the pair is present this way round
m = listX.index(k)
except ValueError: #and if not try its converse
m = listX.index(k2)
return m

def readFile(fname, path): #fills corrgrid by reading list of individual
correlation matrices in input file
IXindiv = [] #index position of patient ID (non-sequentially numbered IDs)
#NameIndiv = 0
f=open(path + fname)
c = 0 #line counter
for line in f:
a=split(line)
try:
t= int(a[0]) #ID number in first cell; it means it's a header row for a
block
##NameIndiv = 0
IXindiv =IXsubjects.index(t) #find the index number for this patient ID
HdrIndiv = a
except ValueError: #first cell is text:it's a data row rather than a
header row
t = 99 #just to give it a value
for i in range(len(a)): #for each column in the row
if i ==0: #first column
RowHdr = a[i][:4] #generate 4 letter row variable name
else: #one of the other columns
ColHdr = HdrIndiv[i][:4] #get 4 letter column variable name from earlier
v = float(a[i]) #get the value of the correlation coeff
if RowHdr != ColHdr: #don't bother if r is 1
CellHdr = RowHdr+' '+ColHdr #gets the new location
listpos = matchXname (IXinteractions,CellHdr) #looks it up in the location function
CorrGrid[IXindiv,listpos] = v #and puts the r in the right place in
CorrGrid

def setMeta(t):
tt=getSet(t)
#print tt
outstring = "Subject,wN,r,lower,upper"+'\n'
sumstring = "nil return"+'\n'
a= zeros((len(tt),12),Float)
ll = len(tt)
if ll<2:
return outstring, sumstring
else:
for i in range(ll):
#a[i,0]=float(tt[i][0]) #ID
a[i,1]=float(tt[i][1]) #N(i)
a[i,2]=tt[i][2] #r(o)(i)
a[i,3]=tt[i][3]**2 #A(i) attrition factor - squared as affects each
data series (Rxx and Ryy)
a[i,4]=a[i,2]/a[i,3] #r(c)(i) 3.29
a[i,5]=a[i,1]*a[i,3]**2 #w(i) 3.32
swi = sum(a[:,5])
mean_rc=sum((a[:,4]*a[:,5]))/swi #mean r(c) 3.33

```



```

mean_ri = sum(a[:,2])/11
for i in range(11):
a[i,6]= (1-(mean_ri**2)**2)/(a[i,1]-1) #var(eo) -of sampling error (i)
a[i,7]=a[i,6]/a[i,3]**2 #ve(i) = var(ec) 3.30
a[i,8]=a[i,5]*(a[i,4]-mean_rc)**2 #weighted squared diff rc from mean
3.34a
a[i,9]=a[i,7]*a[i,5] #weighted ve(i)3.35a
ci=getCI(a[i,2],a[i,1]) #calculate ci for uncorrected r
a[i,10]=ci[0]
a[i,11]=ci[1]
outstring+= tt[i][0]+' '+ str(a[i,5])+' '+ str(a[i,2])+' '+
str(a[i,10])+' '+ str(a[i,11]) +' \n'
var_rc=sum(a[:,8])/swi #variance of corrected r 3.34
ave_ve=sum(a[:,9])/swi #average ve 3.35
var_rho =var_rc-ave_ve
print var_rho
if var_rho>0:
sd_rho=var_rho**0.5
else:
sd_rho =0
credmin=mean_rc-sd_rho*1.96
credmax=mean_rc+sd_rho*1.96
con_r= getCIM(mean_rc,sum(a[:,1]),11,var_rho)
con_rmin = mean_rc-con_r*1.96
con_rmax = mean_rc+con_r*1.96
#het=is_it_het(min(a[:,4]),max(a[:,4]),credmin,credmax)
outstring+= 'pooled,'+ str(sum(a[:,5]))+' '+ str(mean_rc)+' '+
str(con_rmin)+' '+ str(con_rmax) +' \n'
sumstring = str(11)+' '+ str(mean_rc)+' '+ str(con_rmin)+' '+
str(con_rmax)+' '+str(sum(a[:,1]))+' '+ '\n'
outstring+= 'Summary'+', '+str(sum(a[:,1]))+' '+ str(credmin)+' '+
str(credmax)+' ',', '+', '
return outstring, sumstring

def getCIM(r,N,K,v):
if v <0:v=0.001
se = (((1-r**2)**2/(N-K))+(v/K))**0.5 #whitener 3
return se

def getCI(r,n):
zr = 0.5*log((1+r)/(1-r))
sdz=1/(n-3)**0.5
zlo95 = zr-1.959*sdz #95% CIs for z-transformed r
zhi95 = zr+1.959*sdz
lo95=(exp(zlo95*2)-1)/((exp(zlo95*2)+1)) #95%CIs transformed back
hi95=(exp(zhi95*2)-1)/((exp(zhi95*2)+1))
if hi95>1.0:hi95=1.0
return(lo95,hi95)

def OLDsetMeta(i):
dataset = getSet(i)
outstring = "Subject,N,r,lower,upper"+'\n'
sumstring = "nil return"+'\n'
N =len(dataset) #number of subjects with this correlation
if N<2:
return outstring, sumstring

```

```

else:
swr = sn = sw = srd = chisq= 0 #sum of weighted r's,sum of subject n's

for j in range(N):
if RetCode!='y':dataset[j][3]=1.0 #so all subjects have same accuracy
weighting
sn+= dataset[j][1]
sw += dataset[j][1]*(dataset[j][3]**2) #sum of weights
swr+= dataset[j][2]* dataset[j][1]*(dataset[j][3]**2) #sum of weighted r
mwr = swr/sw #mean corrected rc (sum of weight*r / sum of weights
mwr2 = mwr**2
srd=0
for k in range(N): #for each subjects
n = dataset[k][1] #sample size
w = dataset[k][1]*(dataset[k][3]**2) #weight per subject
rc = dataset[k][2]/dataset[k][3] #corrected r = r(original)/attenuation
factor [3.28]
#print "r = " + str(dataset[k][2])+" , rc = "+str(rc)
srd += w*((rc-mwr)**2) #sum of weighted squared differences between rc
and mean rc
chisq+=(w-1)*((rc-mwr)**2)/((1-mwr2)**2)
sdr = (srd/sw)**0.5 #SD of rc's
ser =sdr/(N**0.5)
z = mwr/ser
for h in range(N+1): #each subject + summary line
if h<N:

ro =dataset[h][2]
w = dataset[h][1]*(dataset[h][3]**2) #weighting N*(ICC**2)
r = dataset[h][2]/dataset[h][3] #corrected r (r/ICC)
s = dataset[h][0]#ID
vareo=((1-ro**0.5)**2)/(dataset[h][1]-1) #sampling error variance or robserved
varec =vareo/dataset[h][3]**2 #corrected sampling error variance
svarec +=varec
lo95=(ro- 1.96*sero)/(dataset[h][3]**.05) #correct endpoints by dividing by SQRT of RQ
hi95=(ro+ 1.96*sero)/(dataset[h][3]**.05)
if hi95>1.0:hi95=1.0
outstring+= s+', '+ str(w)+' , '+ str(ro)+' , '+ str(lo95)+' , '+ str(hi95)
+'\n'
else: #for the summary line

n = sn #total weighted N
r = mwr #mean weighted r
s='pooled'
#complicated section for doing z transform of r to set confidence
intervals
sdz= 1/(n-3)**0.5 # standard deviation of z-transformed r
if r >=1.0: r=0.99

zr= 0.5*log((1+r)/(1-r))# ztransformed-r
zlo95 = zr-1.959*sdz #95% CIs for z-transformed r
zhi95 = zr+1.959*sdz
lo95=(exp(zlo95*2)-1)/((exp(zlo95*2)+1)) #95%CIs transformed back
hi95=(exp(zhi95*2)-1)/((exp(zhi95*2)+1))
if hi95>1.0:hi95=1.0

```

```

    outstring+= s+', '+ str(n)+', '+ str(r)+', '+ str(lo95)+', '+ str(hi95)
+ '\n'
    sumstring = str(N)+', '+ str(r)+', '+ str(lo95)+', '+
str(hi95)+', '+str(n)+', '+str(z)+', '+str(chisq)+'\n'
    outstring+= 'Summary'+', '+str(N)+', '+str(z)+', '+str(chisq)+', '
    return outstring, sumstring

```

```

def getSet(j):
    ''' Generates the a pool of matching studies by reading corrgrid and
    identifying all studies with
    the same variable pair (j, an index value from ixinteractions) by looping
    through all subjects(i).
    returns Pairset: a list of tuples of (ID,N,r,ICC)'''
    Pairset =[]
    #print "starting"
    IDtuple=[]
    for i in range(26):
        if CorrGrid[i,j]!=0.0:
            IDtuple.append('S'+str(IXsubjects[i]))
            IDtuple.append(List_of_Ns[i])
            IDtuple.append(CorrGrid[i,j])
            IDtuple.append(List_of_Es[i])
            Pairset.append(IDtuple)
            IDtuple=[]
    return Pairset

```

```

#*****

```

```

OPCode = 'eu' #choose dp if differenced partial; eu if diff(1) pairs; fu if
diff(2) pairs; ru if raw unadjusted"
RetCode = 'y' #code for if this is weighted for test-retest reliability

```

```

p = 'c:/Documents and Settings/Chris/My Documents/_R/MD &
fellowship/Analysis/MetaPairs1/'
f = 'Pooled_corr.txt'
fl = 'subjectN.txt'
if RetCode == 'y':OPCode='C-'+OPCode
else: OPCode='U-'+OPCode
lsl=loadSlists(fl,p)
IXsubjects =lsl[0]
List_of_Ns =lsl[1]
List_of_Es =lsl[2]
IXinteractions = loadXlists()
readFile(f,p)
summary =file(p+"corr_meta_summary_p1.txt",'w',200)
summary.write("Poolcorrs4 run with OPCode = " + OPCode +'at '+strftime("%a, %d
%b %Y %H:%M:%S", gmtime())+'\n')
for i in range(0,150):
    print i
    fname = str(i)+ '-' + IXinteractions[i] #seek out the file with
x= fname.find('/')
fname = fname[:x]+ '~'+ fname[-4:]
#print fname
s = setMeta(i)[1]
if s.find('nil')== -1:

    filestring = p+OPCode+ fname+ '.dat'
    output = file(filestring,'w',30)

```

```
output.write(setMeta(i)[0])
output.close
sumstring = fname + ',' + setMeta(i)[1]
summary.write(sumstring)
#summary.flush
summary.close
```

Appendix I Granger Causality Computer Script (R)

```
folder <- "C:/Documents and Settings/Chris/My Documents/_R/MD &
Fellowship/Analysis/ts/"
filename <- "test"

filename <- "pairs_for_granger"
ext <- ".csv"
library(MSBVAR)
infile <- paste(folder,filename,ext, sep="")
outfile<- paste(folder,"granger4",ext, sep="")
fo = file(outfile,"w")
foheader="causer,causee,Subject, F,P,revF,revP"
cat(foheader,file = fo,sep = '\n',append=T)
maintable<- read.table(infile,header = TRUE,sep = ",")
varnames.vec = colnames(maintable)

for (i in 4:length(varnames.vec)){

  for (j in 4:length(varnames.vec)){

    if(i<j){
      v1<-varnames.vec[i]; v2<-varnames.vec[j];
      s1<-subset(maintable, v1!=999, select = c(Name,i,j));
      s2<-unique(s1[['Name']]); # gets vector of those who are
selected;
      for(k in 1:length(s2)){
        sv <-subset(s1,Name==s2[k],select=2:3);
        if(max(sv)!=999){
          x<- sv[,1]; y<-sv[,2];
          r =granger.test(sv,4) ;
          outstr = paste(v1,v2,s2[k],r[1],r[3],r[2],r[4],sep =
',,');
          cat(outstr,file =fo,sep = '\n',append=T);
          outstr2=paste(v2,v1,s2[k],r[2],r[4],r[1],r[3],sep =
',,');
          cat(outstr2,file =fo,sep = '\n',append=T);
          #note that granger.test regresses y on past y and past x
so report them that way round
          #output is effect of depvar causes testvar
          }
        }
      }
    }
  }
}
close(fo)
```

Appendix J Sample Entropy Computer Script (python)

```
from Numeric import arange, zeros
from Numeric import shape
from Numeric import Float
from Numeric import log
from random import randint
from string import split
from time import gmtime, strftime, sleep

rCount = 0
cCount = 0
FlagList = [1]
inArray = []

def ReadToArray(fname):
#function to return an rCount by cCount array from input file
    BigList = []
    f=open(fname)
    for line in f:
        a = split(line)
        BigList.append(a)

    RBL = len(BigList)
    CBL = len(BigList[0])
    MA = zeros((RBL,CBL), Float)
    for i in range(RBL):
        b = BigList[i]
        #if i == 162:
        #    print b
        for j in range(CBL):

            c = float(b[j])
            MA[i,j] = c
    #print MA
    return MA

def EnCalc(m, r, L, f, mM):
    """
    Function which carries out main sampEn calculation
    m = template size
    r = tolerance, as absolute number derived from proportion of SD
    L = series
    f = whether being done on real or bootstrap series: 1 = real, 0 = surrogate
    mM = number of m+1 matches required
    This is used both by raw data and by surrogates. When no m+1 matches are
found in raw
    data then the flag FlagList[0] remains at 1, if matches found it's set to 0
    Currently there is a threshold of at least 3 m+1 matches in the raw data set
to prevent
    spuriously high readings with 1 or 0 as numerator
    Function returns sample Entropy
    """
    Mmatch = 0 # count of m length matches
    M1match = 0 #count of m+1 length matches
    CList = [] # m+1 length lists
    DList = []
    for x in range(m+1): # load the short lists with zeros
```

```

CList.append(0)
DList.append(0)

for i in range(len(L)-(m)): #i represents start position
    for j in range(m+1): # j is position on comparator list
        CList[j] = L[(i+j)]
    for k in range(len(L)-(m)): #k represents the start of the matching
sequence
        if k!=i: #to avoid self-matches
            for l in range(m+1): #l is position on the matching list
                DList[l] = L[l+k]
            M = MatchUp(CList,DList,r) # make comparisons for m & m+1
            if M == 3: #if m+1 matches, m also does
                Mmatch +=1
                Mlmatch += 1
            elif M == 6: # where only m matches
                Mmatch+=1
            elif M == 0:
                print "error in MatchUp"
# tidy up by deleting short lists
del CList
del DList
#now some coping for nil matches
if Mlmatch== 0: #n m+1 matches
    Mlmatch =.01 # set so that if used for bootstrap will give high reading
    if f == 1: #if raw rather than surrogate
        FlagList[0] = 1 # keep flag at 1 so need to go around again
        print "no m+1 matches when r = " + str(r)
else:
    if f == 1:
        #set threshold mM of m+1 matches in raw data
        if Mlmatch>=mM:
            #print str(Mlmatch) + " m+1 finds when r = " + str(r)
            FlagList[0] = 0
        else:
            FlagList[0] = 1
            # print str(Mlmatch) + " m+1 matches when r = " + str(r) + " fails
to meet threshold of " +str(mM) # so TestnBoot again
            #print str(Mmatch) + " matches at m = " + str(m) + " and " +
str(Mlmatch) + "matches at m = " + str(m+1)
            if Mmatch ==0:
                Mmatch =1
            return -log(float(Mlmatch)/float(Mmatch))

def MatchUp(L1,L2,t):
    """
    Straightforward template comparison function
    Arguments: 2 short lists and t: the tolerance
    Outcomes: returns integer z
    z == 0 startup only
    z == 3 both m and m+1 match
    z == 6 m matches, m+1 doesn't
    z == 9 neither m nor m+1 matches
    """
    x = len(L1)
    z = 0

```



```

if x<> len(L2): #2 line error checking
    print "problem with list length"

for i in range(x-1):
    if abs(L1[i]-L2[i])>t: # one or more items in m length sequence doesn't
match
        z = 9
        break
if z == 0:
    if abs(L1[x-1]-L2[x-1])>t: # m+1 doesn't match
        z = 6
    else: z = 3 #m+1 matches too
return z
"""
End of MatchUp
"""

def stdv(a):
    sx = 0
    for i in a:
        sx += float(i)
    mean = sx/len(a)
    se = 0
    for i in a:
        se += (float(i)-mean)**2
    SD = (se/(len(a)-1))**0.5
    #print 'mean = ' + str(mean)+ '; SD = ' + str(SD)
    return SD
#end of function stdv

def gopherReader(gArray, arrayCols, iters, Case):
    output=[]
    for i in range(arrayCols):
        c = 0
        k=zeros((iters),Float)
        sigsur=0
        x = gArray[0,i]
        for j in range(1,iters+1):
            if x> gArray[j,i]:
                c+=1
                k[j-1]=gArray[j,i]

        surmean=sum(k)/iters
        sursd = (sum((k-surmean)**2)/(iters-1))**0.5
        par_z = (x-surmean)/sursd
        #output.append(str(c))
        print str(Case)+"var"+str(i+1) +',' + str(x) +","+"str(par_z)+ ","
+str(c)+ "," +str(surmean)+ "," +str(sursd)

def superGopher(arrayCols,iters,mReq,Case,Path,m,rz):
    #print' SAMPLE ENTROPY WITH SURROGATE DATA'
    #print 'Medically Unexplained Symptoms Diary Study'
    #print 'Dr Chris Burton'
    #print 'Sample Entropy data for Subject ' + str(Case) + ' (n=
'+str(iters)+' , m = '+ str(m)+ ' , r = '+str(rz)+' )'
    #print strftime("%a, %d %b %Y %H:%M:%S", gmtime())
    a=[]

```

```

ResArray = zeros((iters+1,arrayCols),Float)
for i in range(iters +1):
    if i ==0: #actual data file
        f = Path +str(Case)+'data.txt'
        flag = 1
    else:
        f = Path+str(Case)+'data.csv_surr_%(##)03d '%{"#": i}
        flag = 0
    inArray =ReadToArray(f)
    rCount = inArray.shape[0]
    cCount = inArray.shape[1]
    if cCount!= arrayCols:
        print "00000000000000PS"
    for j in range(cCount):
        for k in range(rCount):
            a.append(inArray[k,j])
            r = rz * stdv(a)
            smpen = EnCalc(m,r,a,flag,mReq)
            ResArray[i,j]=smpen
            a=[]

```

```

gopherReader(ResArray,arrayCols,iters,Case)

```

```

#Here's where it all starts
IDseries=[32,33,35,36,37,38,40,42,44,45,46,48,49,51,52,53,54,55,56,57,58,60,61,63,64,65]
#IDseries=[37,63,64,65]
for CaseID in IDseries:
    Path = 'd:/Tisean/Data/'+ str(CaseID) + '/'
    DataColumns = 8
    Iterations = 500
    Minimum_Matches = 10
    m=2
    r=0.2

```

```

res = superGopher(DataColumns,Iterations,Minimum_Matches,CaseID,Path,m,r)

```

Appendix K Included Papers

Burton C. Beyond Somatisation: a review of the understanding and management of medically unexplained physical symptoms (MUPS). *Br.J Gen.Pract.* 2003;**53**:233-41.

Love T, Burton C. General practice as a complex system: a novel analysis of consultation data. *Fam Pract.* 2005 Jun;**22**(3):347-52.

Beyond somatisation: a review of the understanding and treatment of medically unexplained physical symptoms (MUPS)

Christopher Burton

SUMMARY

Patients commonly present in primary care with symptoms for which no physical pathology can be found. This study is a review of published research on medically unexplained symptoms (MUPS) in primary care. A literature review and qualitative comparison of information was carried out. Four questions were addressed: what is the prevalence of MUPS; to what extent do MUPS overlap with psychiatric disorder; which psychological processes are important in patients with MUPS; and what interventions are beneficial?

Neither somatised mental distress nor somatisation disorders, based on symptom counts, adequately account for most patients seen with MUPS. There is substantial overlap between different symptoms and syndromes, suggesting they have much in common. Patients with MUPS may best be viewed as having complex adaptive systems in which cognitive and physiological processes interact with each other and with their environment. Cognitive behavioural therapy and antidepressant drugs are both effective treatments, but their effects may be greatest when the patient feels empowered by their doctor to tackle their problem.

Keywords: somatisation; medically unexplained symptoms; literature review; qualitative research.

Introduction

A FUNDAMENTAL element of primary care is dealing with symptoms that may, or may not, be due to physical disease. Patients attend with specific symptoms for a variety of reasons,¹ which includes their severity and the disruption they cause, and because of concerns in the patient's mind about what they may represent.² While most people experience at least some physical symptoms, a number of patients repeatedly attend with symptoms for which a conventional pathology cannot be identified. Symptom syndrome³ clusters are widely recognised and include irritable bowel syndrome, chronic pelvic pain, and fibromyalgia. Studies of patients with these conditions have found striking similarities between them,⁴ with a substantial proportion of patients showing evidence of psychological distress⁵ that is either not expressed or is unrecognised in the general practice consultation.

In an attempt to explain this process, psychiatrists have used the term 'somatisation', although the meaning of this term has changed over time.⁶ Initially, it was thought of as being similar to hysterical conversion. Now it effectively has two meanings: the expression of psychological illness through physical symptoms,⁷ (as in the term 'somatised depression'), and repeated medical help-seeking for multiple medical symptoms without organic disease;⁸ for example, in 'somatisation disorder'. These two concepts overlap, but they are not synonymous. To overcome the confusion around the term 'somatisation', many researchers prefer the term 'medically unexplained symptoms (MUPS)'.⁹ While this recognition of uncertainty is helpful in a research environment, the fact that the meaning of physical experiences seems fundamental to these conditions¹⁰ makes it inappropriate for clinical care, and it has been criticised on these grounds.¹¹ With regard to alternatives, 'psychosomatic illness' is seen by the public as synonymous with being 'all in the mind', while 'functional somatic symptoms'⁴ may be preferable, but is not in routine use. In this review, the term 'medically unexplained symptoms (MUPS)' has been used.

Method

Four questions were addressed: what is the prevalence of MUPS? to what extent do MUPS overlap with psychiatric disorder? which psychological processes are important in patients with MUPS? and: what interventions are beneficial? MEDLINE, EMBASE, Cinahl and PsycINFO databases from 1980 to 2001 were searched for any of the following terms: 'medically unexplained symptoms', 'somatization' or 'somatoform disorders', combined with any of the following: 'family practice', 'primary health care' or 'general practice'.

C Burton, MRCP, general practitioner, Sanquhar Health Centre, Dumfries-shire.

Address for correspondence

Dr C Burton, Sanquhar Health Centre, Station Road, Sanquhar, Dumfries-shire DG4 6BT. E-mail: chrisburton@medicine21.com

Submitted: 15 January 2002; Editor's response: 5 March 2002; final acceptance: 29 July 2002.

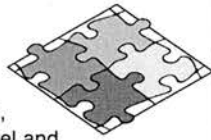
©British Journal of General Practice, 2003, 53, 233-241.

HOW THIS FITS IN*What do we know?*

Medically unexplained physical symptoms, including syndromes such as irritable bowel and fibromyalgia, are common in primary care and are inconsistently associated with mental health problems

What does this paper add?

This paper draws together evidence for complex interactions of physiological and cognitive processes. Neither simple disease syndromes nor a general somatisation disorder are adequate to describe the diversity seen in primary care. Somatisation is too restrictive a label; 'functional somatic symptoms' is a more appropriate term.



An initial list of 570 references was obtained and abstracts from over 300 papers were viewed. Further references were identified from retrieved texts and 137 full texts were obtained and reviewed. These comprised 55 papers from primary care, of which six were about the prevalence of MUPS as the reason for consulting (Table 2), nine were large population datasets (Table 3), eight were intervention studies, and the remainder were interview studies of relatively small groups of patients: 45 studies from non-primary care populations, including five of interventions, and 37 review articles, including five meta-analyses or systematic reviews, which were not repeated for this review. Studies were included in Tables 2 and 3 if they met the criteria of relevance to primary care and included details of numbers of cases. No studies were rejected outright on methodological grounds, although comments on methods appear in the tables. For the wider discussion, where there were no studies from primary care, secondary care or research volunteer studies have been included. Because studies used a wide range of populations, tools, and definitions, formal quantitative comparison was not carried out.

For the purposes of this review, the following definitions were used. 'Medically unexplained symptoms (MUPS)': physical symptoms for which no clear or consistent organic pathology can be demonstrated (although organ dysfunction may be an integral part of the symptom). 'Somatisation': the process by which patients with psychological distress (as measured by psychiatric diagnostic interview or questionnaire) present physical symptoms to their doctors. This process has been further categorised as 'partial somatisation', where the patient acknowledges the possibility that psychological distress may be causing the symptom when directly questioned, and 'true somatisation', where the patient does not acknowledge any psychological link when challenged, despite meeting diagnostic criteria.¹² 'Somatisation disorders': presentation of a specified number of physical symptoms without organic cause in the absence of other major psychiatric diagnosis; the somatoform disorders include DSM-IV Somatization Disorder¹³ (Briquet's syndrome), which is the most severe example. 'Hypochondriasis': a persistent state of increased health anxiety, closely allied to the personality dimension of neuroticism. Because of the pejorative use of the term

'hypochondriac' the term 'heightened health anxiety' has been used in some research.

Results

Figure 1 shows a selection of the symptoms and syndromes under review. Before considering the psychosocial elements of MUPS, it is important to consider recent developments in the pathophysiology of the conditions. Table 1 highlights some of these developments, which demonstrate, first, that current medical knowledge is far from complete and, second, that the boundary between 'organic' and 'functional' may be at least blurred, and at most artificial. Developments in fields such as psychoneuroimmunology¹⁴ are already capable of demonstrating subtle links between physiological processes and emotions.

Studies of the prevalence of MUPS and overlap with psychiatric illness

Studies estimating the prevalence of MUPS in primary care can be grouped into two categories: those that use the main reason for the consultation to determine whether the problem is unexplained or not, and those that apply measures of somatisation to populations that include community samples, primary or secondary care patients, and particular groups such as frequent attenders.

Prevalence of MUPS as the reason for consulting

The search strategy outlined above identified six studies of the prevalence of MUPS as a reason for consulting in primary care (Table 2). The United Kingdom (UK) studies of Mumford¹⁵ and Peveler¹⁶ identified a physical symptom without likely organic disease as the main reason for 15% and 19% of consultations, respectively.

Prevalence of somatisation disorders in primary care and general populations

The search strategy identified nine studies of somatisation disorders, with sample sizes of over 100 individuals from general populations or patients consulting in primary care (Table 3). These used a variety of criteria, but all included patient self-ratings of the presence of symptoms, and used cut-off points based on the number, rather than the character, of symptoms. As well as recording the prevalence of patients reporting above a set number of symptoms, most of these studies identified the prevalence of psychiatric disorder.

The results of these studies are highly dependent on the criteria used both in symptom counts and for severity of psychiatric disorder. While less than 0.5% of patients met the criteria for DSM Somatization Disorder,⁸ which includes at least eight from 40 symptoms owing to non-organic disease in at least four bodily systems, with age of onset before the age of 30 years, 16% to 22% met¹⁷⁻¹⁹ the abridged somatisation criteria of four out of 37 symptoms for men and six out of 41 for women. Over half of one sample of patients²⁰ admitted to at least one MUPS causing some interference with their life. Similar variation in prevalence is seen with concurrent mental illness. While as few as 20% of patients with only one MUPS have a current psychological illness,²⁰ the proportion

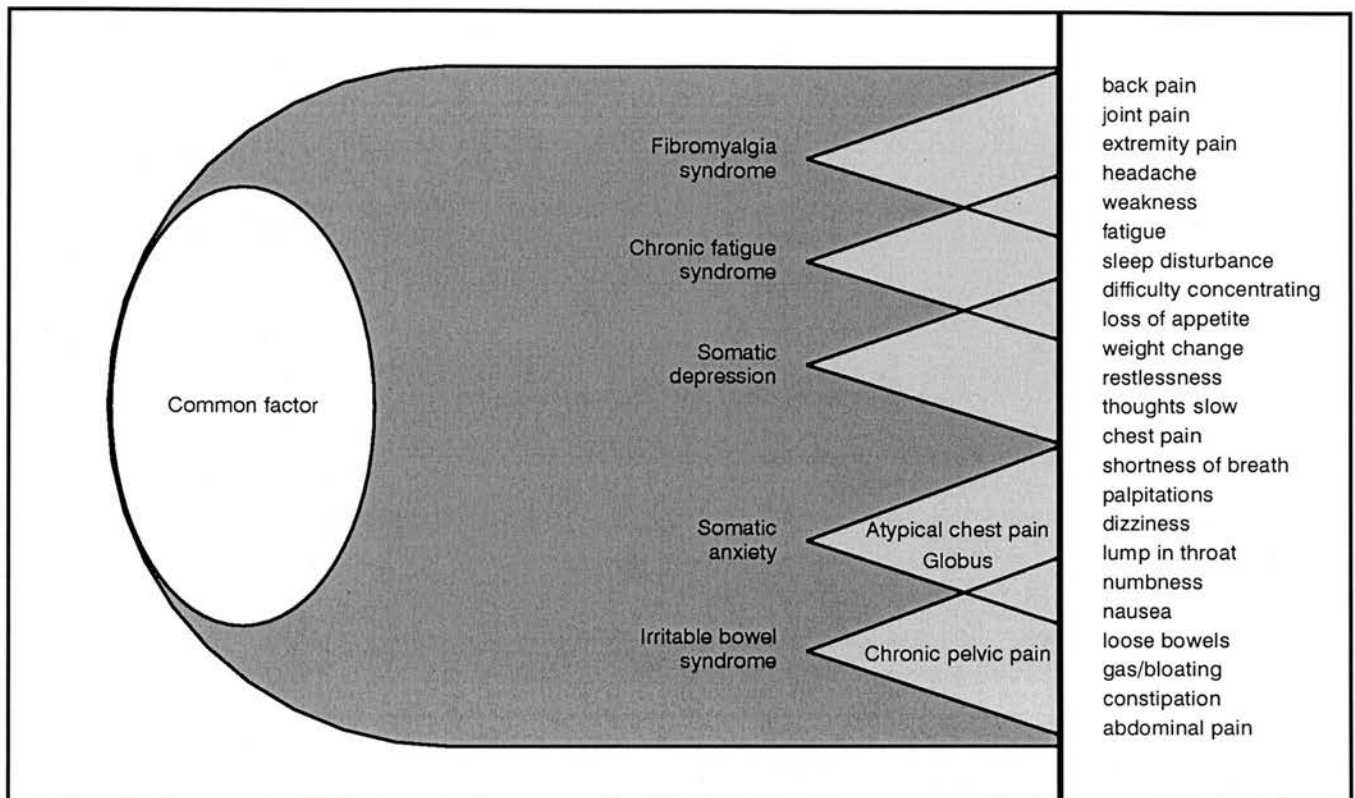


Figure 1. List of functional somatic symptoms showing link to common factor and intermediate syndrome groupings (after Deary²⁴).

Table 1. Recent developments in pathophysiology of MUPS.

Unexplained symptom	Pathophysiological entity	Emerging explanations for some patients
Headache	Neurovascular basis of migraine	Gut neurotransmitters
Dyspepsia	<i>Helicobacter pylori</i> infection	
Dizziness	Benign positional vertigo	Coronary endothelial dysfunction Bacterial overgrowth/fermentation Neural plasticity, excitatory cell death 5-HT neurotransmission
Chest pain/palpitations	Panic disorder	
Irritable bowel		
Chronic pain		
Fibromyalgia		

Table 2. Studies of MUPS as reason for consultation in primary care in Europe and Australasia.

Study	Number in study	Location	Percentage with MUPS	Comments
Mumford 1991 ¹⁵	680 attending for any reason	UK	5 probable, 10 possible	MUPS more likely if past or current depression or anxiety
Peveler 1997 ¹⁶	170 (booked consultations)	UK	19	10% had a mood disorder but presented with physical symptoms, 30% had multiple somatic symptoms, but only one-third of these patients also had a psychiatric disorder
Melville 1987 ⁶⁹	222 (new illness episode)	UK	Not specified at onset, 3 after 6 months	90% of physical symptoms, whether explained or unexplained by organic disease, required no more than two consultations over six months
Palsson 1988 ⁷⁰	78 (booked consultations)	Sweden	16	8/13 with MUPS met hypochondriasis criteria
Pilowsky 1987 ³⁹	100 (booked consultations)	Australia	39	Patients with functional disorders scored higher on scales of affective disturbance and disease conviction
Scicchitano 1996 ⁷¹	112 (new illness episode)	Australia	27	No difference between organic and functional in general health questionnaire score overall. Male patients with functional disorders scored higher on affective disturbance and disease conviction (but $n = 5$). No differences in females

Table 3. Overlap of somatisation and psychiatric disorders in samples of patients from primary care/community care.

Author	Sample size and type	Somatisation measure	Prevalence of conditions (%)	Psychiatric morbidity in patients with unexplained symptoms	Physical symptoms in patients with psychiatric illness
Posse ⁷² 1998	406 consultants, Sweden	Somatisation scale	15 had somatisation on screening	25 had major depression 40 had dysthymia 20 had no psychiatric diagnosis	
Escobar ¹⁷ 1998	1546 consultants, USA	Symptom checklist	22 had f4/37 symptoms (males) or f6/41 symptoms (females)	38% had depression 8% had dysthymia 10% had generalised anxiety 36% had no psychiatric diagnosis	36% had at least one psychiatric illness 13% had psychiatric illness + abridged somatisation 23% had no somatisation
Kisely ¹⁹ 1997	5447 consultants in 14 countries	Reported symptoms categorised as explained or unexplained	6 had >4 (males) or >6 (females) explained symptoms, 13 had >4 (males) or >6 (females) unexplained symptoms, and 2 had >4 (males) or >6 (females) of both	Rates of confirmed psychiatric illness: 4 explained or unexplained symptoms = 10% >4 explained symptoms = 33% >4 unexplained = 69% >4 explained and unexplained = 68% >10 unexplained or 12 explained f80%	52% of those with psychiatric disorder had either >4 explained or >4 unexplained symptoms Social disability was worse if there were unexplained symptoms Physical disability was worse if there were both explained and unexplained symptoms GP detection of psychiatric illness was higher if symptoms were unexplained than if there were none or they were explained
Lobo ⁷³ 1996	1550 consultants, Spain	Symptom count and reason for consulting	27 patients with psychiatric problems with diagnostic interview		9% of population had psychiatric illness but presented with MUPS (over half of diagnoses were of mild anxiety/depression) 20% had <4 symptoms 65% had 4-12 symptoms 15% had >12 symptoms
Kirmayer ^{18,74} 1991	685 consultants, Canada	Symptom count and reason for consulting	11 had major depression or anxiety (n = 75) 26 had one or more of >4 (males) or >6 (females) physical symptoms, hypochondriasis, or somatic presentation of depression	Of patients with MUPS/somatisation, 16% had major depression/anxiety 70% of those with 4-6 functional symptoms had neither major depression/anxiety nor hypochondriasis	Of 75 patients with major depression/anxiety: 17% presented with psychosocial problems 41% presented with physical problems but volunteered a psychosocial basis 23% presented with physical problems but acknowledged the possibility of psychosocial basis 12% presented with physical problems and refuted a psychosocial basis
Munk-Jorgensen et al ⁷⁵ 1997	424 consultants, Scandinavia	'Ill-defined symptoms or mental illness with physical symptoms'	32 had psychiatric illness 17 met somatisation criteria described	66% patients meeting somatisation criteria had psychiatric illness; however, the criteria for somatisation were heavily dependent on presence of psychiatric illness	151 patients with psychological illness: 31% secondary to physical illness 29% presented coincidental physical illness 32% presented somatic symptoms 9% presented purely psychological symptoms

Table 3 (continued). *Overlap of somatisation and psychiatric disorders in samples of patients from primary care/community care.*

Author	Sample size and type	Somatisation measure	Prevalence of conditions (%)	Psychiatric morbidity in patients with unexplained symptoms	Physical symptoms in patients with psychiatric illness
Feder et al ⁷⁶ 2000	172 consultants, USA	'Physical complaints in excess of what would be expected ...'	24 identified as having multiple MUPS by physician 12.5 had current major depression 19 had any current anxiety disorder	Of 42 patients with multiple MUPS: 24% had major depression (versus 9% without MUPS) 48% had any depressive/anxiety disorder (versus 24% without MUPS)	9/21 patients with major depression and 14/33 patients with any anxiety disorder were rated by their physician as having multiple MUPS '... many patients with multiple MUPS acknowledged mental health and social problems'
Fink et al ²⁰ 1999	191 consultants, Denmark	DSM-IV somatoform disorders	54 had one or more (on direct questioning)	30% of patients with one or more MUPS had an additional psychiatric diagnosis Range of 20% for few symptoms of less than 6 months' duration, to 100% for full somatisation disorder	
Weich et al ⁷⁷ 1995	301 consultants, UK	Reason for consulting plus unspecified attribution measure	59 of all patients scored >3 on the general health questionnaire, only 12 of these presented purely psychological symptoms		Approximately 5% of patients presented with purely physical symptoms and attribution, had no evidence of physical illness, and met psychiatric case status on diagnostic interview

rises to over 30% with four symptoms¹⁷ and over 80% with 10 or more,¹⁹ regardless of whether they are medically explained or not.

These studies consistently demonstrate that, while MUPS are common, and often associated with psychiatric morbidity, many patients with MUPS have no definite psychological illness, and patients with multiple symptoms and a refusal to acknowledge a severe mental health problem are rare.

Characterisation of specific MUPS syndromes

Several studies have shown overlaps between the syndromes that comprise MUPS. Many patients with irritable bowel syndrome, for example, meet the criteria for chronic pelvic pain or fibromyalgia, and vice versa. Analysis of population data to identify meaningful classes or disorders have variously suggested 11,²¹ five,²² and four²³ symptom clusters. While such clusters appear broadly to fit clinical patterns, there appear to be no differences between them in terms of psychological characteristics and, indeed, there are many similarities, particularly common aetiological factors and responses to treatment. From this perspective, Deary²⁴ and Wessely⁴ have argued strongly that individual symptoms, while connected to recognised syndrome clusters, are more strongly associated with a single unifying factor, possibly related in some way to the personality trait of neuroticism. Such a three-level relationship is shown in Figure 1.

Aetiological factors

A number of studies have attempted to identify specific aetiological factors for MUPS, although in general the aetiological factors for MUPS are similar to those for anxiety and depression. Deprivation and childhood or family illness^{25,26} may all play a part, as may concurrent stress.²⁷ In women with MUPS there is a higher incidence of past or recent abuse,²⁸ particularly in the case of chronic pelvic pain, in which around a third of patients will have some history of abuse. A longitudinal study of patients at the ages of 36 and 42 years showed that physical symptoms at the first assessment predicted later mental health problems, and also that mental health problems independently predicted future physical symptoms.²⁹

Psychological processes in patients with MUPS

While it seems clear from simple clinical observation that psychological factors are important, defining these has been more difficult, and remains incomplete. Much has been made of the difference between patients with psychiatric illness who present somatically, and those presenting psychosocially.^{12,25} Initial work was grounded in the belief that somatisation represented a flawed process, in which failure to recognise the true problem led to ongoing distress for the patient (and high healthcare costs and frustration for the physician). A systematic review of eight studies of such patients in primary care failed to identify consistent differences between people with psychiatric disorders who present psychologically and those who present with physical symptoms,³⁰ except that, generally, those with physical symptoms were less distressed. Even measures of cognition, such as health anxiety or bodily awareness, appear

little different between 'psychologising' and somatising patients with anxiety and depression. While only around a quarter of affected patients present with purely psychological symptoms, most of the remainder will accept the possibility of a psychosocial component to their physical symptoms, even if they do not volunteer it within the consultation.³¹ Practical reasons, such as lack of time, or a sense that problems are not relevant or amenable to treatment, seem more important than failure to recognise their own mental distress in explaining why patients choose not to disclose psychological problems in consultations.³² Although major depression is harder to recognise if presented somatically,³³ there is conflicting evidence that improving detection improves outcome.

As the terms used for psychiatric illnesses, while undoubtedly important, fail to describe many patients with MUPS, a variety of characteristics and processes have been suggested. While many carry theoretical appeal, some are too broad; for example, 'the personality dimension of neuroticism',²⁴ or too restricted; for example 'alexithymia' — the inability to express emotions in words,³⁴ to be useful in a heterogeneous population. This section examines four processes: hypochondriasis, somato-sensory awareness, attribution (including illness beliefs), and reassurance. To some extent these concepts overlap, and patients who demonstrate one tend to demonstrate others. Nonetheless, they may represent different facets of the problem and warrant further examination.

These processes appear to affect the decision to seek medical attention for any problem, whether 'organic', such as a respiratory infection, or 'functional'. Little and colleagues recently highlighted the importance of MUPS in influencing decisions to consult for all conditions,³⁵ including illness in the subjects' children.³⁶ Studies of these processes, which compare consultations with and without physical disease, may underestimate their importance.

Hypochondriasis

Hypochondriasis is a preoccupation with fears of having, or the idea that one has, a serious disease, based on misinterpretation of bodily symptoms, despite appropriate medical evaluation and reassurance. It overlaps with somatisation, but appears not to be identical; in a study of 184 primary care patients,³⁷ 20% met criteria for hypochondriasis, of whom two-thirds also met somatisation criteria based on the number of symptoms, and a further 20% of the sample met somatisation criteria without hypochondriasis. In another study,³⁸ hypochondriacal patients were more likely to interpret physical symptoms as being due to illness than patients with non-hypochondriacal anxiety, and in two separate studies of healthcare usage hypochondriasis was a predictor of repeated consultation, particularly in men.^{34,39} Robbins and Kirmayer demonstrated hypochondriasis in 10% of over 500 primary care consultants, about half of whom continued to show hypochondriacal beliefs a year later. Improvement in illness worry was matched by improvement in overall well-being, whereas persistence or new occurrence of hypochondriasis was most strongly associated with affective disorder. Hypochondriasis is a common feature of patients referred to secondary care with MUPS,⁴⁰ and it also indi-

cates a greater likelihood of symptoms persisting at follow-up.⁴¹

Somato-sensory awareness

Individuals have varying degrees of bodily awareness. The tendency to notice, and also to amplify, benign sensations is a characteristic found in patients with MUPS. For example, in patients with palpitations but normal investigation results,⁴² high levels of somato-sensory awareness predict persistence of symptoms. The cognitive model of panic disorder, which frequently coexists with MUPS, includes awareness of bodily sensations, which are amplified by the resultant anxiety;⁴³ for example, awareness of heartbeat or breathing triggers arousal, which in turn increases heartbeat or breathing and sets up a cycle. Heightened bodily sensitivity is a feature in many patients with irritable bowel syndrome, fibromyalgia, and chronic fatigue syndrome.

Attribution and illness beliefs

Attribution is the cognitive process whereby somatic sensations are interpreted in the context of the body and its physical and social environment. Using the example of fatigue, attributions can either be normalising ('I'm tired because I'm overworking and unfit'), somatic ('I'm tired because my muscles have been weakened by a virus'), or psychological ('I'm tired because I have depression').

Studies of frequent attenders in primary care⁴⁴ and patients with high health anxiety⁴⁵ suggest that normalising attributions occur less often than in controls. Strikingly, when individuals were asked to write down possible causes for each of 10 common physical symptoms, patients with hypochondriacal anxiety listed an average of eight normalising and 26 psychological or somatic explanations, while non-anxious control patients listed 15 and seven, respectively. The first explanation chosen was that there was normalising 21% of the time in anxious patients, compared to 72% in controls. Frequent attenders were no more likely than patients in the control group to see symptoms as serious, but were less able to come up with reasons why the symptoms might be benign. This may explain why reassurance that rules out problems but does not offer alternative tangible explanations so often fails.

One of the few longitudinal studies of changes in health-related cognitions²⁷ identified a pattern whereby symptoms occurring at a time of newly increased stress tended to be attributed to the stress. Only if the stress persisted did symptoms begin to be presented to doctors as possibly physical. This is compatible with the idea of patients being able to tolerate and normalise symptoms for a limited time before seeking assessment and reassurance that their original attribution was correct.

While doctors have medical models of illnesses, patients also have complex and broadly consistent lay models of health and disease.⁴⁶ Consistent features of these include the name of the condition and its symptoms, the personal consequences of it, how long it will last, and the extent to which it can be controlled or cured.⁴⁷ Patients appear to have health beliefs about individual symptoms as well as established diseases, and Salmon² proposed eight dimensions: four covering aetiology (stress, environment, lifestyle,

and weak constitution), three concerning mechanism (wearing out, internal structure, and internal function), and a final dimension of concern raised by the symptom.

Not only do patients have clear views about their symptoms in their own right, they also view their own experience of the symptoms as at least as important as a doctor's opinion about them. Salmon and colleagues have demonstrated that patients perceive doctors as denying the validity of their symptoms,¹⁰ but that where doctors develop tangible and non-blaming models of conditions with their patients and form constructive alliances against the illness, patients are then able to accept medical opinion.⁴⁸

While MUPS tend to change over time,⁴⁹ attributional style appears to be much more consistent.⁵⁰ Changing specific attributions about symptoms appears to be important in effecting improvement.⁵¹

Reassurance

Illness belief models explore how patients see illness as threatening. Doctors seek to reduce that threat through treatment and reassurance. Unfortunately, reassurance is not always effective; between a third and half of patients report continuing concern about serious illness after normal cardiac ultrasound or angiography.⁵² The effectiveness of reassurance appears to be related to patient characteristics. While all patients who received a normal result after upper gastrointestinal endoscopy⁵³ experienced immediate reassurance, those with the highest levels of health anxiety had returned to original levels of concern within one week, and this persisted for a year.

Psychological models of threat reduction suggest two separate processes:⁵⁴ emotional-heuristic (calming, protecting, and threat-avoiding), and cognitive-systematic (information-seeking and threat-analysing). While emotional, threat-avoiding, reassurance (which may be non-verbal as well as verbal) may be effective in alleviating distress in the short term, it may do nothing to weaken illness representations. If symptoms keep recurring, repeated use of this type of reassurance is likely to produce a cycle of reassurance-seeking and giving that is self-perpetuating.⁵⁵ In contrast, the cognitive model of threat analysing is more threatening in the short term, but more likely to produce long-term changes that in turn can be associated with improvement.⁵¹ Research into minor physical illness suggests that patterns of doctor-patient interaction tend to be self-reinforcing,⁵⁶ and that doctor behaviour in one consultation affects future consultations for the same problem.

Treatment

There have been few studies of treatment of MUPS in primary care. Morris and colleagues⁵⁷ devised a training package to help general practitioners (GPs) recognise depression in patients with MUPS and treat it. The outcomes of a 'before and after' training comparison suggested that patients who acknowledged their depression when it was pointed out to them showed improvements in depression and global function,⁵⁷ and there was a net reduction in healthcare costs.⁵⁸ Agreement between doctor and patient predicted a good outcome, while the patients who denied the possibility of depression did not improve, and felt that their doctors under-

stood them less well after the intervention.⁵⁹

Lidbeck⁶⁰ and colleagues evaluated a programme of group cognitive behaviour therapy (CBT), after thorough physical examination, for patients with MUPS in primary care. Thirty-two subjects were contrasted with 17 waiting-list (eight calendar months) control patients. At the six-month follow-up there were significant changes in illness worry, illness behaviour, and medication usage in the early treatment group, but no change in mood or social problems. No data on subsequent consultation rates are presented. An American randomised controlled trial⁶¹ of a rigidly structured behavioural intervention for patients with MUPS, involving six weekly sessions with homework, demonstrated significant improvements in mood and physical symptoms both one week and six months after the course, compared with waiting list control patients.

There has been only one randomised controlled trial based in primary care of individual CBT,⁶² although a recent systematic review³ identified another 28 studies in secondary care, including over 1600 patients with either mixed unexplained symptoms or specific syndromes. Not all studies demonstrated significant benefit; of those that reported relevant outcomes, CBT improved physical symptoms in 71% of studies, functional status in 47%, and psychological distress in 38%.

Other trials of psychological therapies have generally been small. However, a randomised controlled trial of psychotherapy in 102 patients with irritable bowel syndrome⁶³ showed sustained improvement in symptoms and wellbeing. A recent study of the disclosure of emotionally important events showed no effect on patients' health.⁶⁴

A meta-analysis of antidepressant treatment for MUPS⁶⁵ demonstrated beneficial effects in a wide range of conditions, although not chronic fatigue syndrome. The meta-analysis included 6595 patients in 94 studies (50 of which were of chronic headache). Benefit was seen in 69% of studies, occurring equally in those with or without depression, with an average number needed to treat of three. Because of differences between studies there was insufficient evidence to make detailed recommendations on optimal drugs, doses or duration of treatment.

The importance of a good doctor-patient relationship and of acknowledging patients' concerns has been demonstrated.⁶⁶ Although there is no direct evidence of the effect of consultation behaviour on patients with MUPS, the evidence from a controlled trial that doctor behaviour for minor physical illness affects future consultation rate,⁵⁶ and the observation that a positive, patient-centred approach⁶⁷ improves satisfaction and enablement, and reduces symptom burden and health service usage, point to this being important. In qualitative studies of patients with MUPS, Salmon⁴⁸ identified three types of medical explanation: rejecting (in which patients perceived the doctor as denying the reality of their symptoms, and in which there was unresolved conflict over explanations), collusive (in which the doctor gave in to the patient's interpretation of symptoms but in doing so lost the respect or trust of the patient), and empowering (in which the doctor provided tangible, non-judgemental explanations, which legitimised the patient's suffering and offered opportunities for self management). The empowering

explanations were distinctive, in that patients regarded them as valuable foundations on which to build recovery, or at least cope with their condition in partnership with their doctors.

Conclusion

The notion that most MUPS are the result of a single process of somatisation (particularly the somatisation of mental distress), or are due to a somatisation disorder that can be defined primarily in terms of numbers of symptoms, is no longer supported by the evidence. There is now good evidence that physiology, personality, life experiences, health cognitions, and interaction with healthcare professionals are all important, and any new paradigm needs to include them.

A recent model, which may usefully be explored in understanding MUPS, is that of a complex adaptive system.⁶⁸ In this model the component parts are less important than their many internal and external interactions. Such systems constantly co-evolve with their environment, but tend to organise themselves around states which, while never static for long, are essentially stable. As a result of the dynamic nature of the system, certain properties emerge as a product of the system rather than as a discrete component.

Such a system allows for the kind of complex but inconsistent interactions seen in patients with MUPS, in whom multiple factors interact and illness behaviour patterns evolve within the contexts of the patient's personal life and doctor-patient relationships.

Further research is needed in primary care, particularly in three areas. First, greater understanding is needed of cognitions and the complex way these interact with experiences and symptoms. Such research will draw on qualitative data, but may also exploit longitudinal datasets and models using non-linear analytical techniques. Second, studies are needed of the actual encounters between patients with MUPS, of all levels of severity, and their doctors, to identify and promote the best methods for dealing with these challenging problems. Third, and building on the results in the first two areas, trials are needed to compare enhanced general practice consultations, based on shared explanation and empowerment, as well as re-attribution, with routine care or specialist CBT.

For now the GP's role for patients with MUPS is to validate their experience, provide positive 'empowering explanations'⁴⁸ of symptoms, and to use proven treatments, such as antidepressants and CBT, to modify the process.

References

- Campbell SM, Roland MO. Why do people consult the doctor? *Fam Pract* 1996; **13**: 75-83.
- Salmon P, Woloshynowych M, Valori R. The measurement of beliefs about physical symptoms in English general practice patients. *Soc Sci Med* 1996; **42**: 1561-1567.
- Kroenke K, Swindle R. Cognitive behavioral therapy for somatization and symptom syndromes: a critical review of controlled clinical trials. *Psychother Psychosom* 2000; **69**: 205-215.
- Wessely S, Nimnuan M, Sharpe M. Functional somatic syndromes: one or many. *Lancet* 1999; **354**: 936-939.
- Walker EA, Katon WJ, Jemelka RP, Roy-Bryne PP. Comorbidity of gastrointestinal complaints, depression, and anxiety in the Epidemiologic Catchment Area (ECA) Study. *Am J Med* 1992; **92**: 26S-30S.
- Lipowski ZJ. Somatization: the concept and its clinical application. *Am J Psychiatry* 1988; **145**: 1358-1368.
- Goldberg DP, Bridges K. Somatic presentations of psychiatric illness in primary care setting. *J Psychosom Res* 1988; **32**: 137-144.
- Escobar JI, Gara M, Silver RC, et al. Somatization disorder in primary care. *Br J Psychiatry* 1998; **173**: 262-266.
- Katon WJ, Walker EA. Medically unexplained symptoms in primary care. *J Clin Psychiatry* 1998; **59**(Suppl 20): 15-21.
- Peters S, Stanley I, Rose M, Salmon P. Patients with medically unexplained symptoms: sources of patients' authority and implications for demands on medical care. *Soc Sci Med* 2001; **46**: 559-565.
- Sharpe M, Carson A. 'Unexplained' somatic symptoms, functional syndromes, and somatization: do we need a paradigm shift? *Ann Intern Med* 2001; **134**: 926-930.
- Bridges K, Goldberg D. Somatic presentation of depressive illness in primary care. *J R Coll Gen Pract* 1992; **36**: 9-11.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. Fourth Edition. Washington, DC: APA, 1994.
- Kiecolt-Glaser JK, McGuire L, Robles TF, Glaser R. Psychoneuroimmunology and psychosomatic medicine: back to the future. *Psychosom Med* 2002; **64**:15-28.
- Mumford DB, Devereux TA, Maddy PJ, Johnston JV. Factors leading to the reporting of 'functional' somatic symptoms by general practice attenders. *Br J Gen Pract* 1991; **41**: 454-458.
- Peveler R, Kilkenny L, Kinmonth AL. Medically unexplained physical symptoms in primary care: a comparison of self-report screening questionnaires and clinical opinion. *J Psychosom Res* 1997; **42**: 245-252.
- Escobar JI, Waitzkin H, Silver RC, et al. Abridged somatization: a study in primary care. *Psychosom Med* 1998; **60**: 466-472.
- Kirmayer LJ, Robbins JM. Three forms of somatization in primary care: prevalence, co-occurrence, and sociodemographic characteristics. *J Nerv Ment Dis* 1991; **179**: 647-655.
- Kisely S, Goldberg D, Simon G. A comparison between somatic symptoms with and without clear organic cause: results of an international study. *Psychol Med* 1997; **27**: 1011-1019.
- Fink P, Sorensen L, Engberg M, et al. Somatization in primary care. Prevalence, health care utilization, and general practitioner recognition. *Psychosomatics* 1999; **40**: 330-338.
- Gara MA, Silver RC, Escobar JI, et al. A hierarchical classes analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms. *Psychiatry Res* 1998; **81**: 77-86.
- Robbins JM, Kirmayer LJ, Hemami S. Latent variable models of functional somatic distress. *J Nerv Ment Dis* 1997; **185**: 606-615.
- Liu G, Clark MR, Eaton WW. Structural factor analyses for medically unexplained somatic symptoms of somatization disorder in the Epidemiologic Catchment Area study. *Psychol Med* 1997; **27**: 617-626.
- Deary IJ. A taxonomy of medically unexplained symptoms. *J Psychosom Res* 1999; **47**: 51-59.
- Craig TK, Drake H, Mills K, Boardman AP. The South London Somatisation Study. II. Influence of stressful life events, and secondary gain. *Br J Psychiatry* 1994; **165**: 248-258.
- Hotopf M, Mayou R, Wadsworth M, Wessely S. Childhood risk factors for adults with medically unexplained symptoms: results from a national birth cohort study. *Am J Psychiatry* 1999; **156**: 1796-1800.
- Cameron L, Leventhal EA, Leventhal H. Seeking medical care in response to symptoms and life stress. *Psychosom Med* 1995; **57**: 37-47.
- Fry RP, Crisp AH, Beard RW. Sociopsychological factors in chronic pelvic pain: a review. *J Psychosom Res* 1997; **42**: 1-15.
- Hotopf M, Mayou R, Wadsworth M, Wessely S. Temporal relationships between physical symptoms and psychiatric disorder. Results from a national birth cohort. *Br J Psychiatry* 1998; **173**: 255-261.
- Garcia-Campayo J, Sanz-Carrillo C. A review of the differences between somatizing and psychologizing patients in primary care. *Int J Psychiatry Med* 1999; **29**: 337-345.
- Kirmayer LJ, Robbins JM. Patients who somatize in primary care: a longitudinal study of cognitive and social characteristics. *Psychol Med* 1996; **26**: 937-951.
- Cape J, McCulloch Y. Patients' reasons for not presenting emotional problems in general practice consultations. *Br J Gen Pract* 1999; **49**: 875-879.
- Bower P, West R, Tylee A, Hann M. Symptom attribution and the recognition of psychiatric morbidity. *J Psychosom Res* 2000; **48**: 157-160.
- Jyvasjarvi S, Joukamaa M, Vaisanen E, et al. Alexithymia, hypochondriacal beliefs, and psychological distress among frequent attenders in primary health care. *Compr Psychiatry* 1999; **40**: 292-298.
- Little P, Somerville J, Williamson I, et al. Psychosocial, lifestyle,

- and health status variables in predicting high attendance among adults. *Br J Gen Pract* 2001; **51**: 977-984.
36. Little P, Somerville J, Williamson I, et al. Family influences in a cross-sectional survey of higher child attendance. *Br J Gen Pract* 2001; **51**: 987-994.
 37. Noyes R Jr, Langbehn DR, Happel RL, et al. Health Attitude Survey. A scale for assessing somatizing patients. *Psychosomatics* 1999; **40**: 470-478.
 38. MacLeod AK, Haynes C, Sensky T. Attributions about common bodily sensations: their associations with hypochondriasis and anxiety. *Psychol Med* 1998; **28**(1): 225-228.
 39. Pilowsky I, Smith QP, Katsikitis M. Illness behaviour and general practice utilisation: a prospective study. *J Psychosom Res* 1987; **31**: 177-183.
 40. Speckens AE, Spinhoven P, Sloekers PP, et al. A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. *J Psychosom Res* 1996; **40**: 95-104.
 41. Speckens AE, van Hemert AM, Spinhoven P, Bolk JH. The diagnostic and prognostic significance of the Whitely Index, the Illness Attitude Scales and the Somatosensory Amplification Scale. *Psychol Med* 1996; **26**: 1085-1090.
 42. Barsky AJ, Ahern DK, Bailey ED, Delamater BA. Predictors of persistent palpitations and continued medical utilization. *J Fam Pract* 1996; **42**: 465-472.
 43. Barsky AJ, Barnett MC, Cleary PD. Hypochondriasis and panic disorder. Boundary and overlap. *Arch Gen Psychiatry* 1994; **51**: 918-925.
 44. Sensky T, MacLeod AK, Rigby MF. Causal attributions about common somatic sensations among frequent general practice attenders. *Psychol Med* 1996; **26**: 641-646.
 45. Robbins JM, Kirmayer LJ. Transient and persistent hypochondriacal worry in primary care. *Psychol Med* 1996; **26**: 575-589.
 46. Mabeck CE, Olesen F. Metaphorically transmitted diseases. How do patients embody medical explanations? *Fam Pract* 1997; **14**: 271-278.
 47. Williams S, Weinman J, Dale J, Newman S. Patient expectations: what do primary care patients want from the GP and how far does meeting expectations affect patient satisfaction? *Fam Pract* 1995; **12**: 193-201.
 48. Salmon P, Peters S, Stanley I. Patients' perceptions of medical explanations for somatisation disorders: qualitative analysis. *BMJ* 1999; **318**: 372-376.
 49. Simon GE, Gureje O. Stability of somatization disorder and somatization symptoms among primary care patients. *Arch Gen Psychiatry* 1999; **56**: 90-95.
 50. Garcia-Campayo J, Larrubia J, Lobo A, et al. Attribution in somatizers: stability and relationship to outcome at one-year follow-up. Grupo Morbilidad Psiquica y Psicomatica de Zaragoza (GMPPZ). *Acta Psychiatr Scand* 1997; **95**: 433-438.
 51. van Dulmen AM, Fennis JF, Mookink HG, Bleijenberg G. The relationship between complaint-related cognitions in referred patients with irritable bowel syndrome and subsequent health care seeking behaviour in primary care. *Fam Pract* 1996; **13**: 12-17.
 52. Potts SG, Bass CM. Psychological morbidity in patients with chest pain and normal or near normal coronary arteries: a long-term follow-up study. *Psychol Med* 1995; **25**: 339-347.
 53. Lucock MP, Morley S, White C, Peake MD. Responses of consecutive patients to reassurance after gastroscopy: results of self administered questionnaire survey. *BMJ* 1997; **315**: 572-575.
 54. Coia P, Morley S. Medical reassurance and patients' responses. *J Psychosom Res* 1998; **45**: 377-386.
 55. Neal RD, Heywood PL, Morley S. 'I always seem to be there' — a qualitative study of frequent attenders. *Br J Gen Pract* 2000; **50**: 716-723.
 56. Little P, Gould C, Williamson I, et al. Reattendance and complications in a randomised trial of prescribing strategies for sore throat: the medicalising effect of prescribing antibiotics. *BMJ* 1997; **315**: 350-352.
 57. Morriss RK, Gask L, Ronalds C, et al. Clinical and patient satisfaction outcomes of a new treatment for somatised mental disorder taught to general practitioners. *Br J Gen Pract* 1999; **49**: 263-267.
 58. Morriss R, Gask L, Ronalds C, et al. Cost-effectiveness of a new treatment for somatized mental disorder taught to GPs. *Fam Pract* 1998; **15**: 119-125.
 59. Downes-Grainger E, Morriss R, Gask L, Faragher B. Clinical factors associated with short-term changes in outcome of patients with somatized mental disorder in primary care. *Psychol Med* 1998; **28**: 703-711.
 60. Lidbeck J. Group therapy for somatization disorders in general practice: effectiveness of a short cognitive-behavioural treatment model. *Acta Psychiatr Scand* 1997; **96**: 14-24.
 61. McLeod CC, Budd MA, McClelland DC. Treatment of somatization in primary care. *Gen Hosp Psychiatry* 1997; **19**: 251-258.
 62. Speckens AE, van Hemert AM, Spinhoven P, et al. Cognitive behavioural therapy for medically unexplained symptoms: a randomised controlled trial. *BMJ* 1995; **311**: 1328-1332.
 63. Guthrie E, Creed F, Dawson D, Tomenson B. A randomised controlled trial of psychotherapy in patients with refractory irritable bowel syndrome. *Br J Psychiatry* 1993; **163**: 315-321.
 64. Schilte AF, Portegijs PJM, Blankenstein AH, et al. Randomised controlled trial of disclosure of emotionally important events in somatisation in primary care. *BMJ* 2001; **323**: 86-89.
 65. O'Malley PG, Jackson JL, Santoro J, et al. Antidepressant therapy for unexplained symptoms and symptom syndromes. *J Fam Pract* 1999; **48**: 980-990.
 66. Donovan JL, Blake RD. Qualitative study of interpretation of reassurance among patients attending rheumatology clinics: 'just a touch of arthritis, doctor?'. *BMJ* 2001; **320**: 541-544.
 67. Little P, Everitt H, Williamson I, et al. Observational study of effect of patient centredness and positive approach on outcomes of general practice consultations. *BMJ* 2001; **323**: 908-911.
 68. Cilliers P. *Complexity and postmodernism*. London: Routledge, 1998.
 69. Melville DI. Descriptive clinical research and medically unexplained physical symptoms. *J Psychosom Res* 1987; **31**: 359-365.
 70. Palsson N. Functional somatic symptoms and hypochondriasis among general practice patients: a pilot study. *Acta Psychiatr Scand* 1988; **78**: 191-197.
 71. Scicchitano J, Lovell P, Pearce R, et al. Illness behavior and somatization in general practice. *J Psychosom Res* 1996; **41**: 247-254.
 72. Posse M, Hallstrom T. Depressive disorders among somatizing patients in primary health care. *Acta Psychiatr Scand* 1998; **98**: 187-192.
 73. Lobo A, Garcia-Campayo J, Campos R, et al. Somatisation in primary care in Spain: I. Estimates of prevalence and clinical characteristics. Working Group for the Study of the Psychiatric and Psychosomatic Morbidity in Zaragoza. *Br J Psychiatry* 1996; **168**: 344-348.
 74. Kirmayer LJ, Robbins JM, Dworkind M, Yaffe MJ. Somatization and the recognition of depression and anxiety in primary care. *Am J Psychiatry* 1993; **150**: 734-741.
 75. Munk-Jorgensen P, Fink P, Brevik JI, et al. Psychiatric morbidity in primary public health care: a multicentre investigation. Part II. Hidden morbidity and choice of treatment. *Acta Psychiatr Scand* 1997; **95**: 6-12.
 76. Feder A, Olfson M, Gameroff M, et al. Medically unexplained symptoms in an urban general medicine practice. *Psychosomatics* 2001; **42**: 261-268.
 77. Weich S, Lewis G, Donmall R, Mann A. Somatic presentation of psychiatric morbidity in general practice. *Br J Gen Pract* 1995; **45**: 143-147.

Acknowledgements

I am grateful to Dr Michael Sharpe for discussion of the ideas behind this review and to Dr Phil Wilson and Susie Stewart for their helpful comments on earlier drafts. This work was supported by funding from the Chief Scientist Office of the Scottish Executive Health Department through its Research Practice Scheme.

General practice as a complex system: a novel analysis of consultation data

Tom Love^{a,b} and Chris Burton^c

Love T and Burton C. General practice as a complex system: a novel analysis of consultation data. *Family Practice* 2005; **22**: 347–352.

Background. Complex systems have specific properties of robustness and self organisation which arise from interacting components within the overall system and which govern the system's behaviour. These are typically associated with a power law distribution of event sizes. Commentators have suggested that health systems are complex, but there has been limited quantitative investigation of this issue.

Objectives. To test the hypothesis that consultation patterns in primary care follow a power law distribution typical of a complex system.

Methods. Analysis of 142 050 episodes of non-pathological back pain in routinely collected New Zealand national data. Calculation of the distribution of the duration and number of GP consultations for each illness episode. Secondary analysis of a published UK dataset of consultation rates for 44 000 patients in four general practices.

Results. Number of consultations per episode of back pain demonstrated excellent fit with a power law in the full dataset ($r^2 = 0.96$) and all but one subgroups ($r^2 = 0.90$ – 0.99). The number of consultations per patient from four UK practices was suggestive of a power law distribution ($r^2 = 0.88$ – 0.93).

Conclusions. Consultation patterns in general practice show measurable properties of a complex system. The consistency of the distribution across different population groups suggests that attempts to manage consultation patterns should focus on the whole system of patients, rather than upon individuals or subgroups of the patient population.

Keywords: Behavioural sciences, complex systems, consultations, health service management, primary health care.

Introduction

Complexity and primary care

Complexity science is the generic title for the study of systems whose properties and behaviours arise primarily from the interactions between their individual elements rather than the elements themselves. Complex systems have been observed in a very wide range of settings, from animal biology and ecology, to forest fires traffic jams and the spread of infectious disease.¹

Because the behaviour of the system as a whole results from the interacting individuals within it, patterns across

the system can appear unpredictable. For instance power supply networks (which depend on redistribution of load between interconnected substations) usually adapt well to minor failures in the grid but occasionally a small incident can set off a cascade which overwhelms the system and causes huge areas to lose power. Retrospective analysis of large events typically fails to show any recurring characteristic of their location or timing. Nonetheless at the level of the system, typical statistical properties can be detected.²

Enthusiasm for these ideas from systems theory has led to suggestions that the properties of complex systems might apply to primary care.³ It seems plausible that the large scale patterns seen across primary care as a whole, such as consulting or prescribing distributions, might reflect a complex system. Moreover, the health system is notorious for responding unpredictably to interventions for change. Hospital waiting lists appear to behave as such a complex system.⁴

Patients are known to consult their GP at widely differing rates, only partly explained by differences between

Received 28 October 2004; Accepted 30 December 2004.

^aDepartment of General Practice, Wellington School of Medicine, University of Otago, PO Box 7343, Wellington, New Zealand, ^bHealth Informatics Centre, University of Dundee, Dundee DD2 4AD and ^cCommunity Health Sciences—General Practice, University of Edinburgh, 20 West Richmond Street, Edinburgh EH8 9DX, UK. Correspondence to Tom Love, Health Informatics Centre, University of Dundee, Dundee DD2 4AD, UK; Email: t.love@chs.dundee.ac.uk

individuals in illness, concern about symptoms, and socio-economic factors.⁵⁻⁷ The healthcare system itself influences consultation through accessibility, feedback and cost, while patients' behaviour is influenced by past consultation patterns^{7,8} and their sense of what is appropriate and fair.⁹ The multiple interactions and feedback between individual patients and practitioners, and the healthcare and social systems they comprise, meets the requirements for a theoretical complex system.

Statistical characteristics of complex systems

Analysis of complex systems has identified a characteristic statistical property of complex systems, namely the power law distribution,^{1,10} which relates the magnitude of some value to the frequency of its occurrence within the system. For instance in the case of power networks, the frequency distribution of failures by number of consumers affected shows a characteristic distribution with the great majority of events being very small, but a consistent relationship between frequency and magnitude.

In a power law distribution, the probability p of an event (in this case an episode of illness requiring a particular number of healthcare consultations n) is inversely related to its magnitude n raised to a constant power a , represented as $p \propto 1/n^a$. When plotted on a logarithmic scale, power law distributions form a characteristic straight line with the slope corresponding to the exponent a from the above equation. In addition to suggesting the presence of a complex system, two other features make power law distributions notable in the context of healthcare. Firstly the shape of the distribution is consistent across the entire dataset, suggesting that a common mechanism underlies the effect under study across the range of the data. Secondly, extremely large events (in the case of our hypothesis very high numbers of consultations), while still rare, occur much more frequently in a power law distribution than is the case with an exponential or normal distribution.

We hypothesised that if primary health care constitutes a complex system it should be possible to observe a power law in the distribution of consultation activity. We analysed two datasets: a primary analysis of national data comprising a consecutive series of all patient consultation episodes for back pain beginning in one year, to examine the distributions of duration and number of GP consultations per episode, and a secondary analysis of published consultation rates for patients of four UK general practices.¹¹

We expected that the duration of episodes of illness would be largely determined by factors operating at the level of the individual (such as illness severity and personal obstacles to recovery) but that the number of consultations would be heavily influenced by the system comprising patients, healthcare providers and their respective cultures and social networks and so would show a power law.

Methods

Primary data analysis, New Zealand national back pain data

The Accident Compensation Corporation (ACC) is the statutory funder of accident related health services for the whole of New Zealand. Unlike commercial insurance systems, ACC fund all accident related treatment on a no-fault basis, and without regard to whether lump sum compensation is paid to the injured party. The organisation therefore collects payment data for every item of accident related care in the country, regardless of age, sex or workforce status and includes even minor accidental injuries and strains.

We analysed all episodes of back pain in the New Zealand national accident insurance database which were initiated during 1998. The data included a measure of the number of GP consultations funded under each episode of care and whether the patient had been referred for radiology, physiotherapy or a specialist opinion. Data was stored in Microsoft Access and Excel. Analysis was carried out with these and with custom scripts written in the Python computing language.

For each episode of back pain we estimated the duration, as the time between the first consultation and closure of the claim, and the number of consultations. We then derived frequency distributions for both duration and number of consultations per episode. Data were binned in increasingly spaced groups derived from rounding to the nearest integer of the series $(\sqrt{2})^1, (\sqrt{2})^2 \dots (\sqrt{2})^n$ then converted to logarithms. The resulting data series was plotted and the slope of the regression line and its correlation coefficient r^2 were calculated. The results were tested for sensitivity to different binning schemes. A similar process was carried out for episode duration (measured in weeks and including uncompleted weeks).

The analysis was carried out on the full dataset and by subgroups of patient age and sex. We also used two proxies for episode severity to test their influence on the consultation distribution: whether the episode was managed entirely within primary care or referred and whether the episode resulted in paid time off work.

Secondary analysis of four practice consultation data

To test whether our observations could be replicated in another dataset we took the published data of Neal and colleagues describing consultation rates for four UK general practices.¹¹ This dataset referred to all consultations of 44 146 patients in four practices over a period of 41 months, and was originally a study of the distribution of attendance frequency in UK general practice. We binned the published frequency data into 10 groups by the number of consultations per patient and plots of $\log(\text{consultations per patient})$ versus $\log(\text{number of patients})$ made for each practice and for the pooled data.

Results

New Zealand national back pain data

The dataset contained 148 514 entries. 6564 (4.4%) had insufficient data to calculate duration of claim and number of consultations so were discarded leaving 142 050 episodes of back pain for which at least one GP consultation had been made. 787 episodes of care were still active at the time of data collection in July 2001 but are included in the analysis. The effect of these episodes of care is to make some data points underestimate the number of consultations per claim. These represent 0.55% of the total dataset. Analysis with the 787 episodes excluded made a negligible difference to the results.

Of the 142 050 episodes of back pain, 92 617 (65.2%) involved only a single consultation; 25 084 (17.7%) involved two; 17 533 (12.3%) episodes had three to five consultations; 4717 (3.3%) had six to ten; 1698 (1.2%) had 11 to 20; and 401 (0.3%) had more than 20. Two episodes were for over 90 consultations with the GP. 79 274 (55.7%) episodes were referred outwith the GP surgery: 68 967 to physiotherapy; 22 268 to radiology; 6832 to a hospital specialist and 1492 seen by the GP emergency service).

The consultation count data showed a linear relationship between $\log(\text{consultations per episode})$ and $\log(\text{count of episodes})$ with slope -2.15 ($r^2 = 0.96$, $P < 0.001$) characteristic of a power law relationship. In contrast the episode duration data yielded a skewed gaussian curve which did not fit a power law distribution. Figure 1 shows the distributions of consultation count and duration in both natural (a) and log (b) format. Figure 2 shows that the power law relationship was clearly demonstrated in all subgroups apart from patients who were unable to work. Table 1 contains the results of the subgroup analysis.

Episode duration and consultation count were loosely correlated ($r^2 = 0.24$). When episodes with outlying consultation frequencies were excluded (less than one consultation per 120 days) this correlation became stronger ($r^2 = 0.55$).

UK practice data

Figure 3 shows the individual distributions for the four individual practices and for the pooled data. The correlation coefficient r^2 for the pooled data was 0.91, with individual practices ranged between 0.88 and 0.93.

Discussion

Summary of key findings

The data for consultations per episode of back pain showed an excellent fit with the hypothesised power law distribution; in contrast, the distribution of episode duration did not fit the power law. The UK data for consultations per patient fitted less well, but still

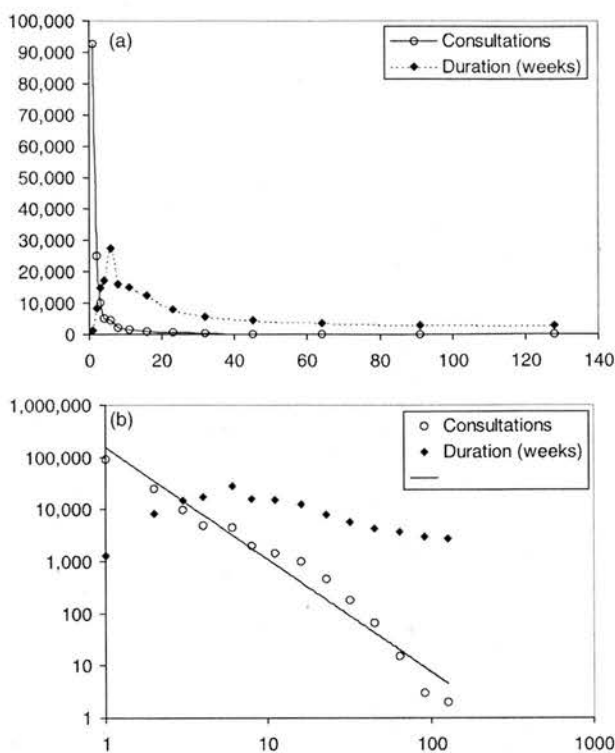


FIGURE 1 Distribution of duration and number of consultations for low back pain shown with conventional (a) and logarithmic (b) axes

approximated to a power law distribution. These results are in keeping with our hypothesis that consultation patterns of individuals generate properties at the level of the healthcare system, as predicted from the theory of complex systems^{1,12} while the duration of each back pain episode was not a function of complex interactions.

Strengths and limitations of the study

The study addresses the knowledge gap between the theory of complex systems and its increasing application in healthcare management. As one property of complex systems is that they can only be fully understood as a whole, the study design runs the risk of artefact due to measuring only one aspect. We addressed that in three ways, by using large and complete samples, by including a measure we predicted would not have a power law distribution (back pain duration), and by corroborating our primary analysis with a secondary analysis of a published dataset.

The unusual nature of the New Zealand ACC, in which GPs are funded on a fee per consultation basis, means that the organisation's data on GP consultations can be easily linked to a specific episode of accident related care covered by the scheme. In the case of musculo-skeletal problems such as back pain, ACC funding covers the medical costs of any episode where there is an acute onset due to an external cause (for instance acute lumbar pain

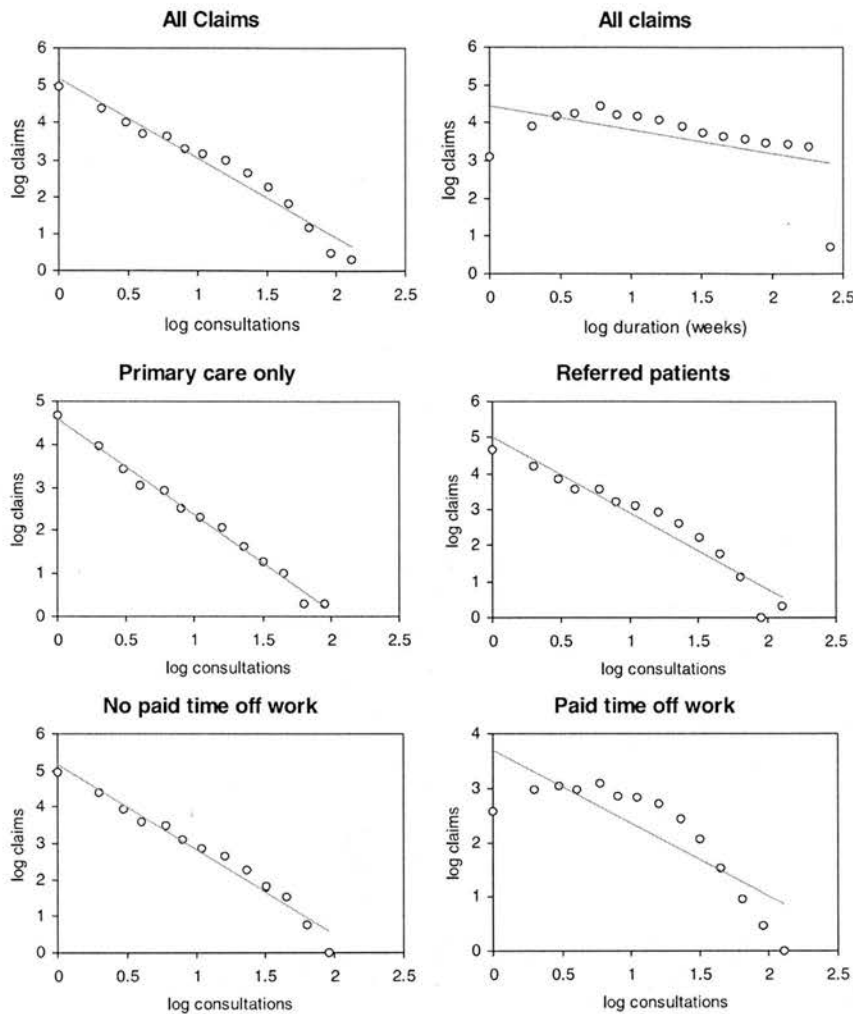


FIGURE 2 Log plots from New Zealand back pain dataset

brought on by lifting one's own furniture would be eligible) and in effect represents all non-pathological back pain.

We considered a number of sources of error in the back pain dataset by excluding incomplete data and testing for the effect of outlying data. We identified 11 989 (8.4%) episodes which averaged less than one GP consultation for every four months of the episode. Of these 7211 had been referred to specialists or physiotherapists and were assumed to be receiving continuing treatment under their care; the remaining 4778 (3.3%) probably represented late completion of claims; these data were retained in the analysis for completeness but recalculation of distributions after their exclusion did not significantly alter the results. 102 (0.07%) episodes involving consultation more often than once every three days were identified, exclusion of these did not significantly change the results.

Both authors analysed the data independently using different methods and software. To reduce bias due to

categorisation, we tested a variety of scales for data binning, to find a series which gave sufficient integers of increasing separation within the range of values. Although the r^2 and slope of the power law distribution varied slightly according to choice of categorisation (by less than 5% and 10% of presented values respectively) the overall pattern was unchanged.

The New Zealand data is collected within the context of a rather unusual primary care funding system which could limit the generalisability of the findings, although we have tried to address this by using UK data to compare the results. It is possible that, in our primary dataset, some of the episodes of back pain represented injuries which included additional factors such as legal compensation to prolong the case. The unique no-fault scheme by which ACC insurance operates minimises this effect. Payment to patients of earnings related compensation when back pain prevented them working could provide an incentive to keep returning to the GP,

thereby distorting the consultation patterns observed under this unusual funding scheme. However such earnings related compensation was paid in only 4.8% of the episodes of care in our dataset, so any distorting effects are likely to be small.

While the data on back pain was for episodes of a single condition, that from Neal and colleagues included

TABLE 1 Details of power law distribution of number of consultations per episode of back pain

	<i>n</i>	Min	Max	slope	<i>r</i> ²
All episodes	142 050	1	100	-2.15	0.96
Primary care only	62 849	1	85	-2.22	0.99
Referred	79 201	1	100	-2.11	0.93
Time off work ^a	6951	1	94	-1.34	0.71
No time off work	135 098	1	100	-2.32	0.97
<i>Age/sex profile</i>					
Male ≤30	24 562	1	94	-2.09	0.97
Female ≤30	18 482	1	40	-1.95	0.98
Male 31-60	43 283	1	100	-2.08	0.95
Female 31-60	36 660	1	85	-1.91	0.97
Male >60	8566	1	64	-1.97	0.98
Female >60	10 497	1	45	-1.85	0.98

Min and Max refer to the respective minimum and maximum number of consultations per episode in the subgroup. Slope and *r*² refer to the characteristics of the regression line of the distribution. *P* values for *r*² are <0.0001 for all categories except time off work.

^a The distribution of this small subgroup did not fit a typical power law curve.

consultations for any reason, including planned follow up of chronic conditions, health promotion, antenatal care and, in some practices nurse and health visitor contacts. Although we assumed the back pain consultations to be primarily patient-led, it is likely that the managed care aspects of general practice, with planned recall for both illness management and health promotion, will have constrained the natural consultation pattern in the UK practice data. We suggest that this explains why the our second data set fits less well to a power law distribution.

Interpretation in the context of existing evidence

There have only been two published demonstrations of the power law distribution in healthcare, both relating to outpatient waiting lists. Nevertheless, power laws have been observed in a wide range of other natural and social settings and can be generated by simulations of complex systems. We suggest that the demonstration of the power law in two separate sets of consultation data provides new evidence to support the notion of primary care as a complex system.

Our analyses add further weight to the argument, supported by the original authors of the UK practice data, that so called 'frequent attenders' are not a discrete group of healthcare users.¹¹ The consistency of the power law distribution across a wide range of consultation behaviour suggests common processes underlying the decision to consult among both high and low consulters. This finding implies that interventions which are intended to manage the distribution of consulting resource across the distribution of patients are *a priori* more likely to be effective if planned across the whole system of consulting patients, rather than targeted at isolated patient groups.

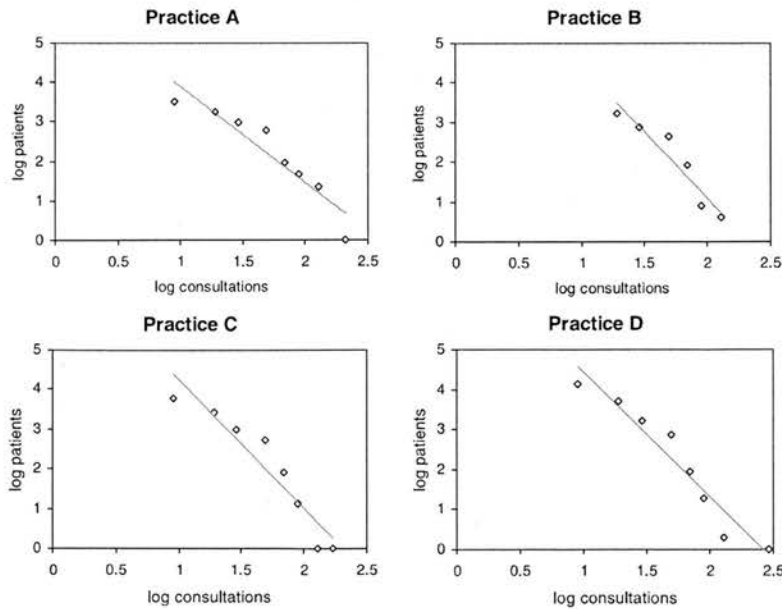


FIGURE 3 Log plots from UK consultation dataset

Implications for research and practice

This study provides the first evidence that, particularly for symptom driven consultations, family practice behaves as a complex adaptive system. This has implications for attempts to understand and shape healthcare. Currently models of consultation rate are based on distributions of predictor variables among independent individuals in a population. Our data suggest that while illness episode duration is distributed at an individual level, use of healthcare is not. Instead, the complex system, comprising patients and their primary healthcare providers, itself strongly influences its own consultation rates.

Such self-organising behaviour is characterised, in experimental models, by an unpredictability in response to stimuli for change. These systems usually respond to change by reconfiguring close to the original state, but occasionally transform: the size of the change often bearing no clear relationship to its trigger. In practical terms, effort to change one part of a system, for instance attempting to address only high users of healthcare, is unlikely to effect long term change as the system will tend to adjust to restore the original distribution. This form of stability has been observed in a number of complex systems including forest fires and traffic flows.¹

More generally, the finding that general practice consultations are a complex system, has broader implications for health services management, particularly in light of the previous observation of power laws in secondary care waiting lists.⁴ At a time when access to primary care in the UK is undergoing major changes and when reductive performance measures are an increasingly common feature of primary care management structures, this study suggests that caution should be exercised in introducing measures which do not recognise the complex nature of primary care, and of health systems more generally.

Conclusion

This is the first study to demonstrate a power law distribution in GP consultation data which is independent of patient characteristics. If consultation patterns in general practice are emergent properties of a self-regulating complex system, then future models for understanding primary healthcare systems need to be capable of explaining this behaviour.

Acknowledgements

The data were provided by the New Zealand Accident Compensation Corporation. We wish to thank Professors Sally Wyke, Tony Dowell and Aziz Sheikh for their constructive comments on earlier drafts of this paper.

Declaration

Funding: CB and TL are supported by the Chief Scientist's Office of the Scottish Executive Health Department.

Ethical approval: the analysis of anonymised patient back pain data was approved by the Wellington Regional Ethics Committee.

Conflicts of interest: none.

References

- 1 Bak P. *How Nature Works*. New York: Oxford University Press; 1998.
- 2 Watts D. *Small worlds: The dynamics of networks between order and randomness*. Princeton: Princeton University Press; 1999.
- 3 Plsek PE, Greenhalgh T. Complexity science: The challenge of complexity in health care. *Br Med J* 2001; **323**: 625–628.
- 4 Smethurst D, Williams H. Are hospital waiting lists self-regulating? *Nature* 2001; **410**: 652–653.
- 5 Campbell SM, Roland MO. Why do people consult the doctor? *Fam Pract* 1996; **13**: 75–83.
- 6 Carr-Hill RA, Rice N, Roland M. Socioeconomic determinants of rates of consultation in general practice based on fourth national morbidity survey of general practices. *Br Med J* 1996; **312**: 1008–1012.
- 7 Jordan K, Ong BN, Croft P. Previous consultation and self reported health status as predictors of future demand for primary care. *Journal of Epidemiology & Community Health* 2003; **57**: 109–113.
- 8 Neal RD, Heywood PL, Morley S. 'I always seem to be there'—a qualitative study of frequent attenders. *Br J Gen Pract* 2000; **50**: 716–723.
- 9 Pollock K, Grime J. Patients' perceptions of entitlement to time in general practice consultations for depression: qualitative study. *Br Med J* 2002; **325**: 687.
- 10 Carlson J, Doyle J. Highly optimized tolerance: A mechanism for power laws in designed systems. *Phys Rev E* 1999; **60**: 1412–1427.
- 11 Neal RD, Heywood PL, Morley S, Clayden AD, Dowell AC. Frequency of patients' consulting in general practice and workload generated by frequent attenders: comparisons between practices. *Br J Gen Pract* 1998; **48**: 895–898.
- 12 Turcotte D, JB R. Self-organised complexity in the physical, biological and social sciences. *Proceedings of the National Academy of sciences* 2002; **99**: 2463.